

RESEARCH AND ANALYSIS

Origins and Evolution of the CASLO Approach in England

The importance of outcomes and mastery when designing vocational and technical qualifications

ofqual

Authors

- Paul E. Newton, Latoya Clarke, Milja Curcin, and Astera Brylka, from Ofqual's Standards, Research, and Analysis Directorate.

With particular thanks to

- Fiona Beedle, Alan Carter, Gillian Cooke, Hayley Dalton, Norman Gealy, Tina Isaacs, Nick Juba, Jill Lanning, Rachael Meech, Judith Norrington, Tim Oates, Barry Smith, Vikki Smith, Isabel Sutcliffe, Jenny Williams, and Chris Winch.

The CASLO Research Programme

This report is part of a series that arose from Ofqual's 2020 to 2024 programme of research into the CASLO approach:

1. The CASLO Research Programme: Overview of research projects conducted between 2020 and 2024.
2. The CASLO Approach: A design template for many vocational and technical qualifications in England.
3. How 'CASLO' Qualifications Work. (This was published in February 2022.)
4. Origins and Evolution of the CASLO Approach in England: The importance of outcomes and mastery when designing vocational and technical qualifications.
5. Responding to Criticisms of the CASLO Approach (Report A): A taxonomy of potential problems.
6. Responding to Criticisms of the CASLO Approach (Report B): Views from awarding organisations.
7. Responding to Criticisms of the CASLO Approach (Report C): Views from qualification stakeholders.
8. Responding to Criticisms of the CASLO Approach (Report D): Properties of qualifications from the CASLO research programme.
9. Understanding Qualification Design: Insights from the 2020 to 2024 CASLO qualification research programme.

Contents

Chapter 1. Introduction	8
The CASLO approach.....	9
Analysis	10
Focus.....	11
Balance.....	12
Chapters	14
Synopsis.....	17
Control	22
Chapter 2. Pre-history	23
1960s.....	24
Training	24
Apprentices	25
Colleges	26
Qualifications.....	27
Problems	34
Summary of the 1960s landscape	43
1970s.....	43
Haslegrave recommendations.....	43
The TEC and the BEC	45
TEC awards.....	45
BEC awards.....	49
Reception	52
City & Guilds	53
Summary of the 1970s landscape	56
1980s.....	59
Roots in North American scholarship	59

Wider sociopolitical context of the 1980s	86
Summary of the pre-history	88
Chapter 3. Genesis	91
NVQs	91
Background.....	91
Design.....	105
Implementation and evolution.....	121
Stocktake.....	141
Conclusion.....	150
GNVQs	151
Design	154
Evolution	160
AVCEs.....	164
Conclusion.....	167
BTECs.....	168
Policies and priorities.....	169
The RVQ, NVQs, and GNVQs	171
Evolving qualification models	174
Evaluations.....	183
Conclusion.....	184
Genesis.....	185
Chapter 4. Dominance	187
The NQF	187
Credit.....	190
Open College Networks.....	191
Further Education Unit.....	192
Policy impetus	194
The QCF.....	195
Reform	195
Transition	197

The framework	197
The CASLO approach	201
Challenges.....	204
Evaluations.....	213
Structure versus quality.....	215
Dominance	218
Chapter 5. CASLO goals	220
Educational goals.....	223
Domain alignment.....	224
Domain mastery.....	225
Qualification efficiency	227
Domain personalisation.....	228
Qualifications and frameworks.....	236
NVQs and GNVQs.....	236
TEC and BEC awards.....	240
OCN awards	241
The QCF	242
Chapter 6. Recent history	244
Post-2010 policy reviews	244
Wolf report	245
Richard report	250
CAVTL (McLoughlin) report	256
Whitehead report	258
VQ Reform Plan.....	261
Sainsbury report.....	262
Policy post-2010.....	264
Post-2010 regulatory decisions	265
Withdrawing the QCF.....	266
Commitment to 'strengthen' VTQs	268
Regulation post-2010	271

Down but not out.....	271
Chapter 7. Conclusions	273
The CASLO approach.....	273
How to understand the approach.....	273
Why the approach was introduced.....	279
How the approach was received.....	283
Its trajectory.....	286
Central control.....	287
Fitness for purpose.....	289
Threats.....	290
Tensions.....	293
Questions.....	294
Lessons.....	302
Reforming TVET qualifications.....	308
Rationalisation.....	308
Control.....	315
Pragmatics.....	318
Conclusion.....	323
References.....	325

Chapter 1. Introduction

There is a lack of organisational memory regarding past policies and programmes in the skills landscape, resulting in an inability to learn lessons.

(City & Guilds, 2014, page 13)

A lack of policy memory in England's Technical and Vocational Education and Training (TVET) sector is widely recognised as a cause for concern (Higham & Yeomans, 2007; City & Guilds, 2014; 2016; 2019; Norris & Adam, 2017; Laczik, Dabbour, Patel, & Wilson, 2023).¹ This is often associated with the frequent churn of officials across government departments and the Civil Service more generally, not to mention the churn of government departments and administrative agencies that occurs over a slightly longer timescale. Churn affects awarding organisations too. All of this churn means that, the longer an assessment system or process has been in place, the harder it will be to locate anyone who understands exactly how or why it was established in the first place, particularly when the period since its introduction is long enough for those originally responsible for its introduction to have retired. This lack of policy memory is compounded by the transience of policy documentation. Although the internet and electronic archives have improved matters, it can still be remarkably hard to locate key policy documents from decades past, especially those produced by administrative agencies that have folded.

Scholarly accounts of the policy landscape help to counteract this lack of policy memory. An important example of relevance to the present report is the book 'Government, Markets and Vocational Qualifications' by Peter Raggatt and Steve Williams (Raggatt & Williams, 1999). But texts like this are few and far between, and they tend to focus more on the sociopolitical context than on the technical detail of qualification and assessment policy.

The present report is different because it is concerned precisely with the technical detail of qualification and assessment policy, focusing specifically on the TVET landscape in England from the 1960s to the present day. Even more specifically, it focuses on a particular approach to TVET qualification design, which became widespread during the 1990s, and which came to dominate the regulated qualifications market during the 2010s. This is the CASLO approach.

¹ Because terminology goes into and out of fashion, we decided to use certain generic terms in ways that might not have been common during the period in question. For example, we use the term 'awarding organisation' throughout, although they were traditionally known as 'awarding bodies'. Similarly, we make frequent reference to 'Technical and Vocational Education and Training', and to 'TVET qualifications', even though we tend nowadays to refer to 'Vocational and Technical Qualifications' (VTQs). When quoting, we use whatever term appeared in the original text.

The CASLO approach

The CASLO approach is a high-level template for designing qualifications – perhaps even a high-level philosophy of qualification design – that is both outcome-based and mastery-based. Outcome-based approaches insist that qualifications ought to be designed on the basis of an authentic and comprehensive specification of the set of learning outcomes that collectively comprises an intended domain of learning. Mastery-based approaches insist that only those students who have demonstrably mastered the full domain of learning should pass the qualification.² We contrast CASLO qualifications with ‘classical’ ones, like GCSEs and A levels, which are neither outcome-based nor mastery-based.³

We have identified 3 core characteristics that are shared by all qualifications within the CASLO family:

1. unit content is specified in terms of learning outcomes (whereas classical qualification content is specified in terms of topics that need to be taught)
2. the unit standard is specified via assessment criteria for each learning outcome (whereas classical qualification standards are holistic, based on mark totals)
3. to pass each unit, a learner must acquire all of the specified learning outcomes, which we refer to as the mastery requirement (whereas classical qualifications do not make requirements concerning specific outcomes)

This also suggests that CASLO qualifications tend to be segmented into units, which is true, although the idea of a single-unit CASLO qualification is entirely legitimate.

As this approach came to dominate the qualifications market in England, it was incorporated into a wide variety of qualifications serving many different purposes in many different contexts. It therefore became an extremely broad family, yet still mainly restricted to the TVET landscape. Until just recently, this family had no distinguishing name. We decided to call them ‘CASLO’ qualifications because they are designed to Confirm the Acquisition of Specified Learning Outcomes (Newton & Lockyer, 2022).

As we shall see, the CASLO approach came to prominence in England during the 1980s and into the 1990s with the introduction of National Vocational Qualifications (NVQs). By the mid-2010s, it had become the default high-level template for

² Outcome-based approaches preceded mastery-based ones, historically, although they often incorporate a mastery-based approach. As such, we sometimes refer simply to ‘outcome-based’ as the superordinate category of most importance to our analysis. We distinguish between outcomes and mastery where necessary.

³ We use the term ‘classical’ to indicate that it is the ‘traditional’ or ‘standard’ approach. We do not mean to imply that it is the ‘definitive’ or ‘highest quality’ approach.

designing regulated TVET qualifications. Although dominant in the landscape by this time, aspects of the approach were criticised in a number of high profile policy reviews. Since then, policy making has tended not to promote the CASLO approach, and in some instances has proscribed it. This raises questions concerning the future of the approach in England.

Unfortunately, there is a deficit of scholarship related to the CASLO approach, which compounds the lack of policy memory and frustrates effective policy making, not to mention effective qualification design, development, and delivery. When NVQs were introduced they attracted a lot of attention from scholars of education. This resulted in a substantial corpus of conceptual critique, but also in a significant body of empirical research. When General National Vocational Qualifications (GNVQs) were introduced a few years later, they also stimulated both conceptual analysis and empirical research related to the approach. Since then, however, there has been little in the way of relevant research and analysis, despite the increasing dominance of the CASLO approach within the TVET landscape.

What this means is that research and analysis have (to date) provided only limited insights into the validity of CASLO qualifications, and into their educational or societal impacts, whether positive or negative. Furthermore, it is unclear how far we can generalise from the insights that we do have – based mainly upon research into NVQs and GNVQs from the 1990s – to regulated CASLO qualifications of the 2020s.

The present report is part of the solution to this deficit of scholarship. It sets out to explain how and why the CASLO approach was first introduced in England, how it came to dominate the TVET landscape, and why it began to fall out of favour with policy makers in recent years. It aims to identify what lessons can be learnt from this trajectory.

Analysis

The work that underpinned this report was predominantly desk-based. This involved studying a variety of documents from the early 20th century to the present day, including:

- articles and books that addressed contemporary debates, as well as some with a more historical bent
- reports, articles, and books that presented original research and evaluation studies
- policy and guidance documents produced by government departments and administrative bodies, including official policy reviews
- publications produced by awarding organisations, including policy statements, communications for centres, and qualification syllabuses or specifications

The aim of the research was to understand the origins and evolution of the CASLO approach in England, within the broader context of training and qualification policies and practices from the 1960s to the present day. This focus helped to make the literature searching and reviewing tractable. Having said that, the story of the CASLO approach is a large part of the story of regulated TVET qualifications in England over the past 3 or 4 decades, so this is a vast literature, impossible to review in its entirety. Triangulating insights from multiple documents was central to the analytical approach, particularly for the period from the 1960s to the 1990s. A few resources proved to be especially useful on particular topics. Where these resources were relied upon heavily, this is made clear in the report.⁴

In trying to make sense of the 1986 Review of Vocational Qualifications, Hargraves (2000) explained that his analytical approach focused on the influence of individual members of the review group. He contrasted his approach with that adopted by Raggatt & Williams (1999), who chose to focus on the influence of key institutions. The focus of the present report is less upon unpacking micro influences on policy formation – whether associated with individuals or institutions – and more upon clarifying the macro rationales that underpinned successive policies and practices, in order to understand the origins and evolution of the CASLO approach in detail.

Our focus on policy is significant to the extent that adoption of the CASLO approach was driven heavily by government, through administrative bodies such as the Technician Education Council, the National Council for Vocational Qualifications, and Ofqual, to name just a few. Although hugely significant in its own right, the approach has always remained just one element of a broader, ongoing policy mission to bring order to what policy makers have often characterised as a disordered ‘jungle’ of TVET qualifications.⁵ So, identifying exactly what the CASLO approach was intended to achieve within this broader policy matrix is challenging, as critical details were not always spelt out in policy documents. Consequently, although the present report is not a defence of TVET qualification policy decisions, it is an attempt to make sense of them, and of how they translated into CASLO qualification practices.

Focus

The story of the CASLO approach intersects with many related stories, to which we cannot do justice within the parameters of the present report. The NVQ model was

⁴ Drafts of this report were also sent out for ‘technical review’ to experts from a variety of backgrounds, many of whom had extensive experience of working for one or more of the key agencies of the 1980s and 1990s (including the NCVQ, the BTEC, the QCA, and so on).

⁵ It is unclear when this intentionally disparaging term was first coined, although it appears, for instance, in the 1986 Review of Vocational Qualifications (De Ville, 1986, page 7).

influential internationally, and versions of it were adopted in New Zealand, Australia, Finland, South Africa, the Persian Gulf, and elsewhere (West, 2004). Outcome-based approaches have also become influential internationally in higher education settings (Cedefop, 2009; Stanley, 2015; Cedefop, 2016). Unfortunately, space and time constraints prevent us from considering either of these important outgrowths.

Even related developments in England that ran in parallel with the introduction of NVQs – concerning Records of Achievement, national curriculum assessment, and GCSE and A level grading standards – have received only limited coverage.

Similarly, our analysis is framed in terms of developments in England, despite those developments often operating in tandem with policies and practices in Wales and Northern Ireland, particularly prior to devolution, and to some extent paralleling developments in Scotland too.

Though we focused purely upon the TVET landscape in England (excluding the higher education sector) we still had to restrict our purview somewhat. Our story revolves specifically around the CASLO approach, which means that important TVET qualification and assessment developments that did not add a great deal to this story have not been discussed. For instance, although we have committed considerable space and time to the rise and fall of the GNVQ, we have not discussed the Diploma qualification – which followed in its wake during the 2000s – as it was not based on the CASLO approach. Likewise, Technical Qualifications that sit within the T Level model also function more like classical qualifications in certain key respects, and therefore receive little attention.

Lastly, it is important to recognise the significance of policy issues that intersected with debates concerning the CASLO approach, but that remained logically distinct, such as the threat to liberal education posed by the ‘new vocationalism’ of the 1980s (Pring, 1995; Stanton, 2012), or the threat to teacher autonomy posed by the shift away from locally developed syllabuses that started during the late 1970s (Bowe & Whitty, 1984), and the establishment of a national curriculum during the late 1980s (McCulloch, 2001). Again, although issues such as these are certainly relevant to debates concerning CASLO qualifications, space and time constraints prevented us from exploring linkages in any depth.

Balance

We have already mentioned the substantial corpus of conceptual critique that emerged during the late-1980s and early-1990s in direct response to the introduction of NVQs, which constitutes a significant chunk of the academic literature on the CASLO approach. This is problematic because, as already noted, it is unclear how far it is possible to generalise from this body of work to current CASLO qualifications.

The NVQ model was an extreme manifestation of the CASLO approach, with a radical approach to specifying learning outcomes (in terms of occupational competence) that took the brunt of academic criticism during the early-1990s. In addition, many scholars saw the introduction of NVQs as an attempt by the government to disempower education providers, which may have further polarised the critique. The passion, if not anger, apparent in many of these early debates has been noted by numerous commentators (see Bates, 1995; Ecclestone, 1997; Hargraves, 2000). Hodkinson reflected:

One of the dangers in the current debate over competence is that the polarisation identified at the beginning of this paper could result in us throwing the baby out with the bathwater.

(Hodkinson, 1992, page 36)

Acutely aware of this polarisation in the literature, we have attempted to be fair-minded in our reconstruction of events, recognising the importance of not throwing babies out with bathwater. As such, we have attempted to do justice to the extent of the problems that beset the introduction of outcome-based qualifications during the 1970s, 1980s, and 1990s, while also explaining why their introduction might still have made a lot of sense to many stakeholders.

Particularly in recent years, there has been a tendency for scholars to treat the NVQ model as though it reflected the essence of all outcome-based approaches. The implication is that, by striking at the heart of the NVQ model, we strike at the heart of the outcome-based approach, per se, and therefore at the heart of the CASLO approach too. On this basis, certain scholars have concluded that outcome-based qualifications are simply not fit for purpose.

Conversely, we think that a more historical analysis reveals that the NVQ model was quite unusual as an outcome-based approach – departing significantly from its ancestors as well as from its descendants – particularly given the uncompromising nature of its early competence model. Rather than embodying the essence of all outcome-based models, NVQs were an idiosyncratic manifestation of what we now describe as the broader CASLO approach, which itself is an idiosyncratic manifestation of the even broader outcome-based approach.⁶ We reflect on whether the problems that became associated with NVQs, GNVQs, and other CASLO qualifications are best understood as inevitable consequences of an unworkable model – as some of the more strident critics have argued – or whether they might alternatively be understood as avoidable consequences from poor implementation. It is fair to say that both NVQs and GNVQs were poorly implemented.

⁶ In other words, NVQs formed a (highly distinctive) subset of the CASLO qualification family, which itself is a subset of the outcome-based qualification family.

The present report goes to considerable lengths to unravel the multiplicity of goals that drove NVQ designers, and designers of other qualifications, to adopt the CASLO approach. This emphasis arose from 2 related observations. First, the CASLO approach was fundamental to the NVQ model, but far from exhausted its innovative design features. So, the NVQ model was driven by goals that went beyond those that drove the CASLO approach specifically. Second, the NVQ model was fundamental to TVET qualification reforms of the 1980s and 1990s, but it was not the only initiative within these reform programmes. We have already mentioned that NVQs were just one part of a broader set of policy initiatives designed to rationalise TVET qualification systems in England. So, it is hard to unpick from the literature – even the literature that focuses specifically on policy goals of the period – exactly which goals the NVQ model, per se, was intended to achieve. Ambiguity of this sort is even more significant for qualifications that were to inherit or adapt the CASLO approach from NVQs during subsequent decades.

We suggest that there are at least 3 distinct perspectives on the goals that drove adoption of the CASLO approach (with distinct goals clustered within each):

1. the certification perspective – to improve the technical quality of assessment (validity)
2. the educational perspective – to improve teaching, learning, uptake, completion, and so on
3. the sociopolitical perspective – to improve the structure of the TVET system

Perhaps understandably, many scholars have focused on unpacking the more subterranean sociopolitical goals underpinning the introduction of outcome-based qualifications (Raggatt & Williams, 1999; Young, 2008). Indeed, some of the more sociological analyses suggest that sociopolitical goals were paramount in explaining their introduction (Young & Allais, 2009; 2011).

The present report attempts to complement analyses of this sort by clarifying goals that ought, in theory, to have been less subterranean – educational goals in particular – but that, in practice, seem not to have been articulated quite as clearly and transparently as they might have been. This attempt to understand rationales for adopting the CASLO approach (and outcome-based approaches more generally) is important because it behoves us to consider whether or not these rationales are just as important today as they might have seemed in previous decades.

Chapters

The story that we will tell revolves around:

- organisations that were responsible for embedding outcome-based qualification models and the CASLO approach more specifically – including the Technician

Education Council (TEC) and the Business Education Council (BEC), which merged to become the Business and Technician Education Council (BTEC), the National Council for Vocational Qualifications (NCVQ), the Qualifications and Curriculum Authority (QCA), and so on

- the qualifications that they introduced – including TEC awards, BEC awards, BTEC awards, National Vocational Qualifications, General National Vocational Qualifications, and so on
- qualification frameworks that embedded the CASLO approach – most notably the NVQ framework and the Qualifications and Credit Framework (QCF)

We identify NVQs, which were introduced during the late-1980s, as the first CASLO qualifications of national prominence. However, TEC and BEC awards, which were introduced during the mid- to late-1970s, were also outcome-based and were critical precursors to the CASLO approach.

We have divided the report into chapters on a roughly chronological basis. Chapter 2 begins by surveying the TVET landscape during the 1960s, in relation to existing qualification systems but also in relation to training more generally, as both halves of this story are important to understanding the origins of the CASLO approach. It also identifies long-standing problems with training and qualifications that help to explain why the approach was adopted. The TEC and the BEC were established during the early-1970s to help solve problems of this sort, and they embedded the concept of outcome-based qualifications (although other awarding organisations were also experimenting with similar ideas). Chapter 2 ends by exploring the wider sociopolitical context of the 1980s prior to the introduction of NVQs, and by tracing the intellectual roots of the CASLO approach back through various North American educational movements.

Chapter 3 explains the crystallisation of the CASLO design template through the introduction of NVQs during the late-1980s. Soon after, it became the design template for GNVQs and for BTEC awards. Chapter 3 explores the design and implementation of each of these 3 qualification types in detail, considering their successes and their failures, as well as how they evolved over time.

Chapter 4 explains how the CASLO approach proliferated during the 2000s (under the National Qualifications Framework) and during the 2010s (under the Qualifications and Credit Framework). While the NQF was positively disposed toward the approach, the QCF made it an accreditation requirement. As QCF qualifications came to dominate the market, it became obvious how dominant the CASLO approach had become. By the mid-2010s, the vast majority of regulated pre-university qualifications in England incorporated the CASLO approach. We explore the route to the QCF by considering the rise of the Credit Movement in England during the 1990s, and the influence of the Open College Network approach.

In Chapter 5, we pause to take stock, and the report becomes more analytical. Reflecting on the story so far, and the documentary analysis that underpinned this project, we attempt to unpick the goals that appear to have driven adoption of the CASLO approach (and outcome-based approaches more generally) over the decades. We identify 4 key educational goals, related to improving:

1. domain alignment – to align curriculum, pedagogy, and assessment as closely as possible with the intended domain of learning (and therefore also with each other)
2. domain mastery – to ensure that all students achieve a satisfactory level of attainment across the full domain of learning
3. qualification efficiency – to make the process of becoming qualified as efficient as possible
4. domain personalisation – to enable the domain of learning to be tailored to the personal situation, interests, or needs of learners (or customised to meet the needs of local employers)

We then explain how these goals appear to have been more or less relevant to the qualifications and qualification frameworks described in Chapters 2 to 4.

Chapter 6 confronts the recent history of the CASLO approach through the lens of 5 key policy reviews commissioned between 2010 and 2015: the Wolf review of 14 to 19 vocational education, the Richard review of apprenticeships, the Commission on Adult Vocational Teaching and Learning, the Whitehead review of adult vocational qualifications, and the Sainsbury review of technical education. While none of these reviews focused specifically on the CASLO approach, each one had something critical to say on the matter. These reviews help to explain why the CASLO approach has fallen out of favour with policy makers in recent years. They precipitated a review of the QCF itself, which was withdrawn in 2015. From that point on, no qualification in England has been required by the regulator to adopt the approach. Where awarding organisations have continued to use it, it has effectively been their choice to do so (although their decision may have been influenced by key stakeholders, such as professional bodies).

Chapter 7 concludes with lessons that we have learnt from our investigation into the origins and evolution of the CASLO approach concerning: the approach itself, its fitness for purpose, and TVET qualification reform more generally. We note that the CASLO approach has tended to be conceptualised and operationalised quite narrowly and rigidly over the years, and we wonder whether it is now time to think about outcome-based and mastery-based approaches more broadly and creatively.

Synopsis

This penultimate introductory section aims to explain exactly what this report is all about – the CASLO approach – and to explain exactly what we mean by it. Exactitude is important, here, because the level of analysis that we have adopted – the design template level – is quite unusual. It is far more common for research projects to be pitched at the level of a particular qualification (A level physics, for example) or at the level of a particular qualification type (General National Vocational Qualification, for example). The present analysis is pitched at a much higher level of analysis because it concerns core characteristics that, at the outset of our research programme, were common across a very large number of very widely divergent qualification types. Part of the mystery that we wanted to solve was why this happened to be the case.

CASLO qualifications are easy to spot because the heart of each CASLO unit specification tends to look like this (albeit with variable numbers of learning outcomes and assessment criteria):

Learning Outcomes	Assessment Criteria
(The learner will...)	(The learner can...)
LO 1 [xxx]	AC 1.1 [xxx]
	AC 1.2 [xxx]
LO 2 [xxx]	AC 2.1 [xxx]
	AC 2.2 [xxx]
LO 3 [xxx]	AC 3.1 [xxx]
	AC 3.2 [xxx]

As such, the substance of each CASLO unit specification comprises a list of learning outcomes, which set out what a learner will need to be able to do, to know, to understand (or suchlike) in order to pass the qualification. And, for each learning outcome, a corresponding set of assessment criteria explains the basis for confirming that each learning outcome has been acquired, in terms of what the learner will need to have demonstrated. To be awarded a CASLO qualification, a learner will need to have achieved all specified learning outcomes across all relevant units, and this will typically require them to have satisfied all of the criteria for all of the outcomes.

We introduced the concept of a CASLO qualification back in 2022 to allow us to refer to the very many qualifications that we regulated that were based on this high-level design template, which stipulates:

1. tightly specified outcomes
2. tightly specified criteria
3. a stringently applied mastery principle

This (and nothing more) encapsulates what we mean by the CASLO approach. It is true that these core characteristics are often associated with other design features, including a substantial amount of college-based or work-based assessment. But it is important to emphasise that CASLO qualifications differ widely in terms of their stated purposes, design features, and implementation processes. Again, this was part of the mystery that we wanted to solve – why exactly the same design template came to underpin a plethora of very differently conceived qualifications.

On one level, the mystery proved not to be very mysterious at all. During the late 1980s, the CASLO approach was specified as an accreditation criterion for any qualification that was to be part of the NVQ framework. At that time, it was anticipated that all technical and vocational qualifications would end up being accredited to this framework, and many were.⁷ Furthermore, during the late 2000s, the CASLO approach was specified as an accreditation criterion for any qualification that was to be part of the QCF. Once again, it was anticipated that all regulated technical and vocational qualifications would end up being accredited to this framework, and the vast majority were.⁸ So, in a very important sense, the proliferation of the CASLO approach in England is directly attributable to the intervention of government-sponsored agencies and their accreditation requirements, in the broader context of perennial efforts to rationalise the TVET qualification landscape.

Yet, on another level, there is a much more interesting story to tell. This includes different agencies adopting very different positions on qualification design, despite fundamentally agreeing on the significance of outcomes and mastery when designing vocational and technical qualifications. It also includes how awarding organisations responded to the opportunities and threats associated with the imposition of the CASLO approach. It is clear that some awarding organisations coped better than others, and some coped quite poorly. Others found creative ways to align long-standing, well respected practices to the CASLO approach, despite the approach being unsuited to their circumstances.

⁷ This process was incentivised in many instances by rules stipulating that only accredited qualifications would be eligible for funding.

⁸ This was also incentivised by qualification funding rules.

As noted above, we concluded that NVQs were the first CASLO qualifications of national prominence. Yet, we accept that this claim could be debated. First, it could be argued that NVQs were not truly CASLO qualifications because they were specified in terms of ‘elements of competence’ (not learning outcomes) and ‘performance criteria’ (not assessment criteria). More fundamentally, they were defined independently of any particular course of learning, which raises the question of whether elements of competence and outcomes from learning are really the same kind of thing. In response, we would argue that they are, in fact, extremely similar, both structurally and functionally, and it is precisely this structure and function that is fundamental to the CASLO approach.⁹

Second, it could be argued that NVQs were not actually the first CASLO qualifications of national significance. It could instead be argued that the first generation of BTEC qualifications, which were rolled out during the mid-1980s, deserve this title (or perhaps even the TEC or BEC awards that preceded them). In response, we would argue that the distinctive template of tightly specified learning outcomes and assessment criteria, alongside a stringently applied mastery principle, is far more clear-cut in the NVQ model (and in subsequent BTEC models) than in qualifications that were rolled out during the 1980s by the Business and Technician Education Council. That said, the early TEC, BEC, and BTEC awards were explicitly outcome-based and also mastery-based, so they clearly prefigured the CASLO approach in many important respects, even though we argue that they are best seen as critical precursors.

There is another reason for describing NVQs as the first CASLO qualifications of national prominence, which relates specifically to their public profile. NVQs and GNVQs were widely criticised, which included criticism of their underlying model, which embodied the CASLO approach. Many academic journal articles contributed to this critique, and many book chapters too. Yet, it also played out in the public domain, which included a high profile, highly critical Channel 4 Dispatches TV programme fronted by Professor Alan Smithers of the University of Manchester:

Few are aware of the new revolution transforming education for the majority in Britain or the unconventional approach now being adopted. Even fewer are aware that many involved in this revolution, often by circumstance rather than desire, are expressing grave reservations (see Box B). They fear the new system, far from raising the profile and establishing the credibility of vocationally-based education, may discredit it further.

(Smithers, 1993, page 10)

⁹ We also note that elements of competence were actually described as “outcomes of learning” when NVQ criteria and procedures were first specified (NCVQ, 1988, page 9; NCVQ, 1989, page 3).

This report described how these ‘revolutionary’ new qualifications – based on elements of competence and performance criteria – were displacing well known and respected qualifications, such as BTEC Nationals and City & Guilds craft certificates. It also gave the impression that both City & Guilds and the BTEC were essentially opposed to these radically new qualifications. As such, the report helped to construct a public narrative for NVQs and GNVQs (and the underlying CASLO approach) which cast them as fundamentally different from existing qualifications and fundamentally disliked by established awarding organisations. Consequently, the prominence that NVQs and GNVQs acquired during the early-1990s bathed the CASLO approach in a fairly negative light.

There is one final reason for describing NVQs as the first CASLO qualifications of national prominence, which relates to the idea of a design template. Whereas the TEC, the BEC, and later the BTEC, worked in partnership with colleges, enabling them to offer rich and integrated teaching, learning, and assessment programmes with national currency, the National Council for Vocational Qualifications operated at one step removed from this, specifying the high-level approach that awarding organisations needed to adopt when developing qualifications. This entrenched the idea of deferring to a centrally-specified design template as the basis for building a qualification.

Because of the controversy that surrounded the introduction of NVQs and GNVQs, we have discussed their introduction, rollout, and reception in considerable detail. This helps us to separate criticisms that might be specific to NVQs, and criticisms that might be specific to GNVQs, from more generalisable criticisms of the CASLO approach. However, we have also emphasised just how important it is to understand what came before NVQs and GNVQs, because they were just part of a growing zeitgeist of enthusiasm for outcome-based and mastery-based qualification models. Contrary to the impression given by critics like Smithers, both City & Guilds and the BTEC were in the vanguard of this revolution. They certainly had issues with how NVQs operationalised the CASLO approach. But both organisations acknowledged the critical role that outcomes, criteria, and mastery ought to play when designing vocational and technical qualifications, having pioneered their use for years.

To conclude this section, our research and analysis led us to a variety of conclusions concerning the approach, which included:

- the CASLO approach is a high-level template for designing qualifications, which stipulates tightly specified outcomes, tightly specified criteria, and a stringently applied mastery principle
- during the late-1980s, the CASLO approach crystallised within the original NVQ model, and soon after within the original GNVQ model

- the roots of the CASLO approach can be traced further back, however, as qualification bodies increasingly embraced the critical role of outcomes, criteria, and mastery throughout the 1970s, building these features into a variety of vocational and technical qualifications
- accreditation criteria for the NVQ framework, and subsequently the QCF, resulted in the CASLO approach achieving almost hegemonic status by the mid-2010s

We also concluded that:

- as a high-level template, the CASLO approach only fixes a few core design features, and therefore provides little more than the foundation for a fully elaborated design template (which would be bespoke to any particular qualification type) – as such, different types of CASLO qualification might well differ significantly in terms of their validity
- the CASLO qualification family includes qualification types that have succeeded, qualification types that have failed, and qualification types that should never have been required to adopt the approach – in short, the approach is neither universally fit for purpose nor universally unfit for purpose
- the CASLO approach is not sacrosanct – despite having achieved almost hegemonic status, there are other ways of designing outcome-based qualifications and other ways of designing mastery-based qualifications – and this invites us to think creatively about the significance of outcomes and mastery when designing vocational and technical qualifications for the future

What follows is the story of the CASLO approach: its pre-history, its genesis, its dominance, its goals, its fall from favour with policy makers, and reflections on its possible future. We end by exploring broader issues related to reforming TVET qualifications, which became apparent from our study of the CASLO approach. The most important of these issues concerned the risk of conceptualising and operationalising qualification reform too narrowly, with insufficient attention to the wider education and training changes that are necessary for a reform to bed in, particularly the need to support teacher and trainer development from the outset. Qualification reforms are best understood as education and training reforms that are initiated through changes to certification requirements. When reforms focus squarely on assessment, with little attention to the wider context, they tend to fail. In other words, if we invest in developing high quality qualifications – but fail to invest simultaneously in developing high quality teaching and training – then we cannot expect to reap the rewards we seek.

Control

Before jumping into the pre-history of the CASLO approach, it is worth noting a critical issue that runs throughout this account, which relates to how control has been exercised over time. Although Ofqual, which came into force in 2010, was the first 'official' qualification regulator in England, the control of TVET qualification systems has become increasingly centralised over time, particularly since the 1970s, and this is key to understanding the CASLO story.

The introduction of the TEC and the BEC during the 1970s was an important watershed. These government-sponsored bodies were established to help co-ordinate the provision of technician and business qualifications, and to confront the problem of unco-ordinated proliferation of courses. Although they did help to co-ordinate and rationalise the TVET qualification landscape, they did not actually regulate it. Up to the late-1980s, the major players – which included the BTEC, City & Guilds, the RSA, and others – remained essentially autonomous, self-regulating bodies.

The NCVQ was established during the late-1980s to develop and populate the NVQ framework. This (and subsequent) framework(s) substantially increased central control over the TVET qualification landscape via accreditation requirements – criteria that any qualification would need to satisfy in order to be accredited to the framework – in tandem with centrally specified course funding requirements that heavily favoured accredited qualifications.¹⁰ Note that these arrangements made it possible for government to exercise control over the design of TVET qualifications, thereby assuming responsibility for decisions that would previously have remained firmly under the control of the major players.¹¹ This helps to explain how the CASLO approach came to dominate the landscape.

¹⁰ In previous years, courses and qualifications tended to be locally funded (largely by Local Education Authorities) so this lever had only recently become available to central government.

¹¹ It is worth emphasising that the NCVQ exercised control at a higher level than the TEC, BEC, or BTEC. The Technician-Business councils functioned much like awarding organisations, although they were officially constituted as validating bodies (engaging primarily with centres). The NCVQ was an accrediting organisation, which made it more like a regulator, with a remit to influence systems and practices across the qualifications sector (engaging primarily with awarding organisations).

Chapter 2. Pre-history

The National Vocational Qualification (NVQ) framework was developed during the late-1980s as a mechanism for ‘rationalising’ the so-called ‘jungle’ of Technical and Vocational Education and Training courses and providers. We will argue that NVQs were the first CASLO qualifications of national prominence in England. To understand why the NVQ framework was introduced – and to understand the genesis of the CASLO approach more generally – we will need to appreciate the state of TVET provision during the 1960s and 1970s, and the role that TVET qualifications played within it.

The TVET training landscape changed significantly during the 1960s, as did the TVET qualifications landscape during the 1970s. Subsequently, the relationship between training and qualifications evolved significantly during the late 1980s, with the introduction of NVQs. This zeitgeist of change reflected increasing recognition that neither work-based (on-the-job) training systems nor college-based (off-the-job) qualification systems were functioning adequately, and that both needed to be fixed.¹²

Although training in England had traditionally been seen as a matter for industry and commerce, it had become clear by the end of the 1950s that this voluntaristic system – based largely on agreements between employer associations and trade unions (Hansen, 1967) – was not working, and that government would need to intervene to orchestrate a national solution (Sheldrake & Vickerstaff, 1987). Intervention in training was most evident during the 1960s as the 1964 Industrial Training Act rolled out. Intervention in qualifications was most evident during the 1970s as the Technician Education Council and the Business Education Council were established, and during the 1980s as the National Council for Vocational Qualifications was established.

Just as today, a distinction was drawn between professional qualifications and technical and vocational ones. Professional groups such as engineers, accountants, physicians, and so on – often with Royal Charters – had well-established entry requirements and strong links to universities that provided high-level qualifications

¹² During the 1960s and 1970s, there was a clear distinction between ‘further education’ (which happened in colleges) and ‘training for skill’ (which happened at work, although sometimes also in colleges, albeit somewhat differently). As explained by Cantor & Roberts (1972), the purpose of further education was to provide the underpinning knowledge and understanding required for successful job performance, as well as to enable employees to cope with change, or to support progression to more advanced study. Conversely, the purpose of training for skill was to develop the skills required to perform a job competently in situ. This formal separation between education and training in England can be traced back to the 1870s (Hansen, 1967).

tailored to those requirements. The qualifications at the heart of the present report, however, are the technical and vocational ones that were not provided by universities.

1960s

The following subsections describe the most prominent pre-university TVET qualifications of the 1960s, explaining how they were delivered and assessed. To help understand the origins of the CASLO approach, particular attention will be paid to problems that beset these qualifications. We will begin with a short discussion of the training context during the early 1960s, to which we will return at the end of the 1970s section because it constitutes a critical piece of the CASLO jigsaw puzzle.

Training

During the 1960s, many young people left school and entered the workforce at the age of 16, some having successfully completed General Certificate of Education Ordinary levels (GCE O levels) and some not. Many young people left school even earlier, at 15, typically with no formal qualifications.

Cantor & Roberts estimated that only a small proportion of school leavers entered some form of skills training during the early 1960s, perhaps fewer than 20% of boys and 3% of girls (Cantor & Roberts, 1972). Many trainees received no off-the-job training, and even those formally designated as apprentices often received inadequate training, as many employers saw apprenticeship as little more than a source of cheap labour, and as small firms were reluctant to invest in apprentices for fear of them leaving for bigger firms once qualified. In short, industrial training: “was, to a large extent, obsolete and out-of-date” and “failing to produce the goods” (Cantor & Roberts, 1972, page 81).

These problems were recognised during the late-1950s and early-1960s within a series of policy reviews (including the Carr report: Carr, 1958) and policy statements (including the ‘Industrial Training’ white paper: Ministry of Labour, 1962). In response to this pattern of unco-ordinated and reluctant training provision by individual companies, the 1962 white paper identified 3 objectives for a national training policy (Sheldrake & Vickerstaff, 1987):

1. to link training provision to wider economic and technological needs and developments
2. to improve standards of training provision
3. to spread the costs of training more evenly across companies

The subsequent 'Industrial Training Act' was enacted in March 1964, establishing a Central Training Council and associated Industrial Training Boards (ITBs). By the end of 1970, 27 ITBs had been created, representing industries as diverse as engineering, shipbuilding, agriculture, printing, catering, and knitting (Cantor & Roberts, 1972). The Boards comprised employers, trade union representatives, and educationists. Under the Act, the ITBs were required to:

- publish recommendations on the nature, content, and length of training suitable in their industry, including associated further education
- ensure that adequate facilities were available for the training required

Given the nature of the levy system set up to support this remit, Cantor & Roberts suggested that the function of the ITBs was, in effect, to persuade industry to train its employees better (Cantor & Roberts, 1972). This remit embraced industries with an established tradition of training (including engineering) and industries that did not have this tradition (including construction).

Although it has been criticised, and judged by some to have failed (see, for example, Sheldrake & Vickerstaff, 1987), the Act signalled a critical change of direction for intervention in training, and facilitated an increase in the quality and quantity of industrial training (Cantor & Roberts, 1972; Wheatley, 1976; Huddleston & Unwin, 2024). In subsequent sections, we will consider how the work of the ITBs related to the origins of the CASLO approach.

Apprentices

During the 1960s and 1970s, apprenticeship was the normal means of training for employment, based upon a more or less formal contract with an employer. Apprenticeship was the mechanism by which employers in many occupations exercised their generally accepted responsibility for training employees.¹³

Wheatley (1976) identified a variety of different forms of apprenticeship in England:

- craft apprenticeships – for skilled manual occupations
- technician (or technical) apprenticeships – for technician level occupations in industry
- commercial apprenticeships – similar to technician apprenticeships, but in commerce rather than industry

¹³ The following analysis of the nature of apprenticeship during the 1970s borrows heavily from a detailed European Commission report by Wheatley (1976).

- student apprenticeships – for students at university or professional institution level, linked to sandwich courses (to qualify for membership of a professional institution)
- graduate apprenticeships – for holders of a university degree or comparable qualification who were training in an industrial, commercial, or professional field (to qualify for membership of a professional institution)

A very large majority of apprenticeships were taken up as first employment by 16- and 17-year-old school leavers, and craft apprenticeships dominated the landscape. The apprenticeship model involved a combination of on-the-job training (provided by employers) and off-the-job training (provided by further education colleges). Off-the-job training mainly involved apprentices attending part-time day courses or block-release courses during working hours without loss of pay.

Colleges

Although the college system had already been ‘rationalised’ by the Ministry of Education by the 1960s, it was still possible to identify at least 5 broad categories: Colleges of Advanced Technology, Regional Colleges, National Colleges for specific industries, Area Colleges, and Local Colleges. These were in addition to Evening Institutes and English as a Foreign Language establishments (Pedley, 1964).

Those who attended college for off-the-job training tended to be classified into 1 of 4 categories, which reflected the ‘grade’ of their job – operative, craftsman, technician, or technologist (Peters, 1967).

Operatives were semiskilled workers who carried out specific operations using machinery or plant. Aspiring operatives would typically have left school at 15 and would generally not be employed as apprentices. They might be following a college course, but not necessarily one that led to a nationally recognised qualification.

Craftsmen were manual workers who carried out skilled practical tasks (for example, an engineering fitter, or a maintenance electrician).¹⁴ Aspiring craftsmen would often be apprentices, attending college on a day-release scheme to study for a craft qualification.

Technicians were specialist assistants to technologists, requiring not only practical aptitude but also a good knowledge of relevant mathematics and science (for

¹⁴ We have decided to use the gendered term ‘craftsmen’ in this section of the report because this was the term that was used in the 1960s. The vast majority of learners studying for City & Guilds awards for craftsmen, or for National awards for technicians, were boys or men. Pedley (1964) provided related statistics for 1961. Just over 28,000 girls or women were studying for City & Guilds awards, compared to just over 317,000 boys or men. Just over 4,000 girls or women were studying for the Ordinary National Certificate, compared to just over 144,000 boys or men.

example, an assistant designer, an instrument artificer, or a skilled lab worker). Aspiring technicians would typically be apprentices. After having successfully completed a 5-year course of secondary education, they would attend college on a day-release, sandwich, or full-time scheme to study for a National award.

Technologists would have studied the fundamental principles of their subjects. They would be capable of initiating change, accepting a high degree of responsibility, and potentially pushing forward the boundaries of knowledge (for example, a university graduate in engineering). Aspiring technologists would typically be studying for a university degree, a Diploma in Technology, or a related qualification in a university or College of Advanced Technology.

Qualifications

Very many different kinds of qualification were offered during the 1960s.¹⁵ In 1959, D.E. Wheatley, who was a Deputy Director at City & Guilds, published an authoritative overview of its qualifications, which he classified into 12 categories:

1. Plant Operative
2. Plant Operative, Higher Grade
3. Junior Craftsman
4. Average Craftsman
5. Craftsman, Higher Grade
6. Technician
7. Technician, Higher Grade
8. Technologist
9. Threshold of Management
10. Extension Subject
11. Teacher's Certificate
12. Domestic Subject

Craft and technician qualifications were the pillars of TVET provision during the 1960s, particularly the Craft Certificates awarded by City & Guilds and the Ordinary and Higher Nationals awarded by Joint Committees. We will consider these 2 qualification suites shortly.

¹⁵ In 1966, City & Guilds offered exams in 282 subjects, from 19 subject groups. The subject groups included Mechanical Engineering, Mining & Quarrying, Vehicles, Textiles, Building, Distributive Trades, and so on (Peters, 1967).

Entry to craft and technical courses was heavily triaged – to ensure that all students achieved as much success as their talents and inclinations permitted – with different pathways open to students entering after 4 versus 5 years of secondary education, preparatory courses where necessary, and ample opportunity for transfer across courses during the early stages (Pedley, 1964). Students embarking on the craft route after 5 years of secondary education might study for 3 years (part-time) before taking their Intermediate Certificate exams. If successful, they might then study for a further 2 years (part-time) before taking their Final Certificate exams. Similarly, students embarking on the technician route after 5 years of secondary education (and gaining 4 GCE O level passes) might study for 2 years (part-time) before taking their Ordinary National Certificate exams.¹⁶ If successful, they might then study for a further 2 years (part-time) before taking their Higher National Certificate exams.

Major players

Although a large number of awarding organisations offered qualifications of this sort, some were more influential and dominant than others. City & Guilds of London Institute (hereafter City & Guilds) was the most important awarding organisation of this period, although other major players included:

- the Regional Examining Unions (REUs) later known as the Regional Examining Boards (REBs)
- the Royal Society of Arts (RSA)
- the London Chamber of Commerce (LCC)
- Pitman's Institute

A book entitled 'A Parent's Guide to Examinations' by F.H. Pedley (1964) is an excellent source of information on the range of TVET qualifications available during the 1960s, some of the most prominent qualification suites being the:

- City & Guilds Intermediate, Final, or Advanced Certificate
- City & Guilds Full Technological Certificate
- Ordinary National Certificate (and Diploma)
- Higher National Certificate (and Diploma)
- Diploma in Technology

¹⁶ If entering after 5 years of secondary education, but without the exemption granted by 4 GCE O level passes, they would generally be expected to study a General course for a year. This would have been a normal route into ONC or OND study during the mid-1960s.

Pedley distinguished technical exams from non-technical ones, including exams in commerce and art. This resonates with the distinction between industry (technical) and commerce (business) that was later to structure recommendations from the Haslegrave committee, which we will consider shortly. Some of the most prominent non-technical exams included the

- Certificate in Office Studies
- Certificate in Business Administration
- Diploma in Management Studies
- Diploma in Art and Design

Even today, it is notoriously hard to describe the TVET qualification landscape, which has evolved to cater for a wide variety of learners, with a wide range of learning needs, in a wide variety of learning contexts, across a wide range of sectors. The landscape of the 1960s was even more disparate and even harder to describe. Having said that, we can get a sense of the lie of the land by considering some of the most prominent TVET qualifications of the time and the learners that they catered for. Specifically, we will focus upon the:

- very wide range of awards for craftsmen provided by City & Guilds, and the
- more restricted range of national awards for technicians, including the National Certificates and National Diplomas at Ordinary and Higher level

Together, these qualification suites constituted the major pillars of college-based TVET provision in England during the 1960s. They were also the focus of much debate. The following sections describe these 2 major qualification suites and the debates that engulfed them.

Craft Certificates

During the 1960s, City & Guilds was the principal awarding organisation for craft qualifications. It was best known for its Intermediate Certificates and Final Certificates – which corresponded to basic and advanced craft syllabuses – although it also offered full Technological Certificates, which were at technician level, as well as lower-level courses for operatives.¹⁷ Cantor & Roberts described City & Guilds courses as the “meat” or the “staple diet” for the average student in further education (Cantor & Roberts, 1972, page 67).

¹⁷ Pedley noted that the Intermediate Certificate had recently been relaunched as the ‘Craft Certificate’ and the Final Certificate had recently been relaunched as the ‘Advanced Craft Certificate’ (Pedley, 1964, page 141).

City & Guilds qualifications were designed to respond to the distinct needs of each industry or setting, which meant that different courses might specify different entry standards, different course lengths, different numbers of grades, different amounts of practical work, and so on. Accordingly, there was no such thing as a 'typical' City & Guilds craft qualification (Wheatley, 1959; Cantor & Roberts, 1972).

City & Guilds differentiated between 'basic' courses (typically lasting 3 years) and 'advanced' courses (extending to 5). Learners who wished to follow an advanced course would first need to pass the Intermediate Certificate, before being allowed to work towards their Final Certificate. Peters (1967) noted that most craft courses were based on a time allowance of one day plus one evening per week, which amounted to around 240 hours per year (8 hours x 10 weeks x 3 terms).

Although City & Guilds awarded the Final Certificate in craft – and was the only organisation to do so – end of year exams leading up to the Final Certificate tended to be arranged by individual colleges or (where they existed) by Regional Examining Unions (REUs). To ensure consistency of approach, the colleges and REUs liaised with City & Guilds over entry requirements, course content, and exam standards. In fact, only a minority of craft apprentices progressed to the Final Certificate. Most finished their training at Intermediate Certificate level, after 4 or 5 years of study, to become competent 'journeymen'.

City & Guilds believed that its primary purpose in holding exams was to promote the establishment of courses of study appropriate to the needs of industry. Wheatley proposed that the:

objective is to design a course in which the average student will find interest and stimulation and be able to make steady progress, so that he will be beneficially exposed to educational influences throughout the course and profit on both the technical and the general educational sides.

(Wheatley, 1959, page 41)

City & Guilds relied heavily upon advisory committees for guidance on preparing course syllabuses. These committees were widely representative of the education service (at national and regional level), employer associations, trade unions, and ITBs (Wheatley, 1976). They provided a forum for co-operation between industry and the education service in defining, monitoring, and developing further education courses. Syllabus content typically included: craft theory, practical workshop or laboratory activities, allied subjects, industrial studies, and general studies. Pedley noted that syllabuses were available for one shilling, providing colleges with details of content, expectations, recommended text books, teacher guidance, and so on (Pedley, 1964).

On assessment, Wheatley (1959) noted the importance of assessing practical ability. Practical tests were provided in addition to written exams. For instance, the Heating

and Ventilating Operatives' Practical Course (course 179) culminated in one written paper of 3 hours' duration and one practical test of 5 hours' duration. Pedley provided further insight into the assessment of craft courses by reproducing in full the 2 written papers from a City & Guilds (Intermediate) Craft Certificate in Plumbers' Work. The first 2-hour paper comprised 25 questions, from which candidates were to answer 20. The first 3 questions read:

1. State the normal height:
 - (a) to the top front edge of a kitchen sink in a sink unit;
 - (b) to the top front edge of a pedestal wash basin;
 - (c) of a W.C. pan.
2. If a rectangle is 4 ft by 3 ft, calculate the length of the diagonal.
3. What is the difference between a "separate" and a "combined" system of underground drainage?

(Pedley, 1964, page 164)

For the second 2-hour paper, candidates were provided with a drawing board, a sheet of paper, and logarithmic tables. They were required to answer 5 questions from 8, and the first 2 read:

1. Describe, with the aid of sketches, the action of an automatic flushing cistern for use on a range of urinals.
2. Define FOUR of the following terms: static head, vacuum, latent heat of fusion of ice, induced siphonage, specific gravity, maximum density of water.

(Pedley, 1964, page 166)

An historical account of City & Guilds prepared (largely) by a former Secretary to the Institute, Peter Stevens, provides additional insight into its approach to grading craft exams (Stevens, 1993). He noted that, following the introduction of computer processing in 1968, a new grading scale had been adopted for City & Guilds certificates: passed with distinction, passed with credit, passed, or failed. Performance on individual papers was also recorded on a scale from grade 1 (high) to grade 8 (low), with grades 7 and 8 counting as fail grades.

It is worth noting the attention paid to tailoring syllabuses and assessments to the differing needs of different learner groups. For instance, when designing courses for low-level operatives (for example, boiler operatives) Wheatley noted that special attention should be given to:

- (a) the provision of a course of limited duration (*i.e.* with the objective clearly visible ahead) and with a character and tempo attuned to the interests and

capacity of unselected secondary modern school leavers and men who may have been away from any form of education for many years; [...]

- (d) strict limitation of mathematics and science; that which is included must have immediately obvious relevance to practical operation and usually be presented as an integral part of the technical syllabus;
- (e) examination papers – these should not call for lengthy written answers nor deal with matters outside the candidate's experience or responsibility, though they do not by any means require to be all of the 'yes/no' type;

(Wheatley, 1959, pages 38 to 9)

Ordinary and Higher Nationals

Following the first world war, a system of National awards was created on an industry-by-industry basis in response to perceived inadequacies of traditional 'Science and Art' technical exams (Foden, 1951). During the 1920s, arrangements were confirmed for the Institution of Mechanical Engineers, the Institute of Chemistry, the Electrical Engineering Institution, and the Institute of Gas Engineers. Schemes for other industries followed in the 1930s and 1940s. National awards in Business Studies were introduced during the early-1960s to replace earlier schemes that had not been successful, including Commerce. By the mid-1960s, there were around 15 National awards or similar schemes (Foden, 1966).

Most schemes awarded 4 distinct types of qualification:

1. Ordinary National Certificate (ONC) – roughly A level standard
2. Ordinary National Diploma (OND) – roughly A level standard
3. Higher National Certificate (HNC) – roughly pass degree standard
4. Higher National Diploma (HND) – roughly pass degree standard

Certificate and Diploma courses were based upon similar syllabuses, although the Certificate route was only open to part-time students and the Diploma route was only open to full-time students. The Diploma route therefore provided for a much broader treatment of the learning domain. Certificate courses required at least 2 years of study and at least 150 hours of study per year. This typically meant studying at least one afternoon and 2 evenings per week, or one day and one evening (Peters, 1967).¹⁸

¹⁸ Accounts seem to vary in detail. For instance, Haslegrave observed that: "At the time the white paper was written, the minimum specified time for technician courses was 180 hours in each year, although 220 was more usual for part-time students" (Haslegrave, 1969, page 15).

By 1969, around 3,000 of the approximately 5,600 candidates that entered HND exams were in engineering subjects, with the next largest entry being for Business Studies, with around 1,900 candidates. From 1960 to 1969, numbers of full-time HND entrants had grown substantially, more than 5-fold, while numbers of part-time HNC entrants – although slightly higher in 1969 (around 18,000) than in 1960 (around 16,000) – were actually declining from a high (of nearly 22,000) in 1968. Cantor & Roberts explained that this decline was likely to continue due to a change in policy regarding the currency of HNC exams with professional bodies (Cantor & Roberts, 1972). Similar patterns of uptake were evident for ONDs and ONCs.

National awards were administered under the authority of a Joint Committee that represented the relevant professional institution(s), the Ministry of Education, and teacher associations. For instance, National awards in Business Studies were made by the Joint Committee for National Awards in Business Studies, with representation from the major professional institutions in advertising, accounting, banking, secretarial work, building societies, sales managers, purchasing officers, and other bodies including the Ministry of Education (Pedley, 1964). Each Joint Committee was responsible for entry requirements, syllabuses, and national standards. Business Studies exams, however, were developed and delivered jointly by the RSA and the LCC.

For other industries, it was often REUs or individual colleges that developed and delivered the exams. So, for most schemes, this was a system of local development and delivery, linked to national entry requirements, content, and standards. National standards were ensured via moderation:

The examinations are set and marked by college teachers or (commonly in smaller colleges) by a regional union, but they are assessed by an external examiner appointed by the professional institution concerned and acting under the control of the joint committee. Homework, drawings, notebooks and so on may be called for by the assessors.

(Peters, 1967, page 98)

Montgomery noted that end-of-year exams were governed largely by individual colleges (or REUs where they operated) although external assessors were appointed by the Joint Committees for the final exams (Montgomery, 1965). These assessors scrutinised draft exam papers before they were administered and scrutinised the marks that were awarded after exams had been sat.¹⁹ Assessors might also request sight of classwork, including notebooks, drawings, and suchlike. Montgomery explained that this had the beauty of devolving considerable

¹⁹ Montgomery noted that up to 40% of questions might be replaced by the assessors, and up to 40% might be made compulsory.

responsibility for assessment processes to colleges, which enabled local interests to be catered for while also giving the awards a national currency.

Although written exam papers were a very important part of the assessment process, other evidence was also taken into account for each award:

It is a most important feature of all technical college examination work that *homework, class work and practical work* are taken into account, as well as *satisfactory attendance*. Indeed in all these aspects the college must be satisfied before entering a candidate. For example, in the case of O.N.C. courses in engineering, certificates are awarded only to those who (a) have passed the examination in all subjects in each year of the course; (b) have made at least 60 per cent of the total possible attendances in each year and in each subject; (c) have obtained at least 40 per cent in homework, laboratory work and drawings separately in each subject and in each year of the course; and (d) have reached an overall average of 50 per cent of the marks.

(Pedley, 1964, pages 154 to 5)

Different accounts of the awarding process emphasise different details. For example, Montgomery noted that:

Students had to score at least 40 per cent of the total possible number of marks in each subject at the 'finals', and do likewise for homework in the last year. Course work and examination marks were to count towards the ultimate total in the ratio of 30 per cent to 70 percent. 50 per cent of the total possible number of marks were to earn a pass, 85 per cent would win a distinction. Such were the arrangements in a typical scheme, but it should be borne in mind that the system was extremely flexible, and catered for different courses in different parts of the country.

(Montgomery, 1965, page 215)

The written exam format was highly valued. For example, the Business Studies committee was particularly insistent that even ONC students should demonstrate "logical thought and correctness in writing" and required them to produce an extended essay of 3 to 5 thousand words, prepared under the guidance of their tutor, in addition to their final exams (Pedley, 1964, page 174).

Problems

The radical changes that occurred during the 1970s and 1980s – which included the introduction of an outcome-based NVQ model that emphasised workplace learning and assessment – were a response to earlier problems, some of which had come to a head by the end of the 1960s. The following discussion of challenges facing TVET qualifications during the 1960s is selective, but it identifies certain key issues, and a

changing zeitgeist, which resulted in the development of outcome-based approaches to qualification design.

Crowther report

In March 1956, the Central Advisory Council for Education (England) was asked to advise the Minister of Education on the education of boys and girls between 15 and 18. Chaired by Sir Geoffrey Crowther, the committee reported 3 years later (Crowther, 1959), addressing a wide range of issues, which included:

- why the school leaving age should be raised from 15 to 16 (which eventually became law in 1972)
- the case for a lower-level examination below the O level (which was to become the Certificate of Secondary Education)
- the sixth-form and problems of university entrance

Of particular relevance to the present report, it also discussed the neglected educational territory of around a quarter of the national cohort of boys and girls who finished school at 15 or 16, but who continued to spend a significant part of their time in further education, training, or instruction. These were typically technical apprentices and trainees who studied part-time in technical colleges. The report identified 2 great challenges for this sector, which the committee believed ought to be solved in tandem:

1. many more skilled craftsmen and technicians were required to support the needs of industry and agriculture (via an alternative route to the grammar schools)
2. there were far too few young people who stayed in full-time education from 16 to 18 (only 1 in 8)

The committee expressed particular concern over the effectiveness of the part-time courses that lay at the heart of provision for craftsmen and technicians. First, 270 hours per year (which they associated with one day release and one evening class per week) was insufficient to cover essential ground for the technical exams, let alone to enable additional studies. This tended to focus teaching narrowly on preparation for exams. The committee argued that aspiring technicians, in particular, should spend more time in college obtaining a deeper and broader education, ideally on 'sandwich' courses (Crowther, 1959).

Second, the committee was worried about failure and drop out rates in technical education, particularly for ONC and Intermediate Crafts courses, both of which (during the late-1950s) were 3-stage courses. The report provided evidence on the percentages of students successfully completing each stage, which were roughly similar for both courses. Only around two-thirds of those who started an ONC or Intermediate course completed their first stage successfully. Only around a half

completed their second stage successfully. Less than a third completed their final stage successfully and were awarded a certificate.

Third, they were worried about 'retardation' problems beyond non-completion. The stages just mentioned were intended to correspond to a year of study, but "a good many students spend a good many years" on a single stage (Crowther, 1959, page 360). The report noted that, of those who were eventually successful on the 5-stage HNC, 19% had spent 7 or more years studying for it.

Finally, the committee believed that the answer to re-engaging boys and girls "who lose their intellectual curiosity before they have exhausted their capacity to learn" lay in an alternative approach to knowledge, a more 'practical' approach (Crowther, 1959, page 391). The idea of 'practical' education was not very popular during the late-1950s, but the committee wished to rehabilitate it. Education should include both practical and theoretical elements. Yet, young people of 15 or 16 ought to be given a choice between a predominantly 'academic' route or a predominantly 'practical' one. The committee ended by identifying 2 as-yet-unsolved problems:

First, how is the programme of practical work to be designed so that the intellectual stimulus and the theoretical knowledge arise out of it? We suspect that too often, even when both elements are present, they remain separate. Secondly, how can the practical work and the intellectual value deriving from it best be assessed? Any education in England which aims at equipping its pupils for a professional status has to conform to an examination system designed in relation to an educational curriculum of which both main subjects and the approach to them are academic. It will be apparent that it is not always easy to reconcile this parallel road with traditional examinations. Sometimes it can only be done with undesirable distortion. Some of the most valuable aspects of the education it can give would, we suggest, more naturally be tested by a scrutiny of work done during the course and by an oral examination upon it.

(Crowther, 1959, page 399)

Wastage, retardation, failure

The Crowther report highlighted very serious problems for TVET qualification systems that revolved around Craft Certificates and National awards. Many students dropped out before taking their exams. Others took their exams, failed, and then dropped out. Yet others spent many years repeating exams before succeeding. These problems were certainly not new. Indeed, Foden had discussed them in 1951, suggesting that many students on such courses were simply not suited to them, often repeating their experiences of failing many times before ultimately giving up. The problem, he suggested, was that students were not being selected effectively (Foden, 1951). The Crowther committee observed that problems of failure were

exacerbated by the requirement to pass National awards in all subjects each year before being allowed to progress (Crowther, 1959). Taylor & Beaumont (1967) echoed this concern, noting that this was why O levels had been transformed into single-subject awards, in contrast to the School Certificate grouped award that preceded them.

The 1961 white paper 'Better Opportunities in Technical Education' (Ministry of Education, 1961) set out to address problems of 'wastage' (drop out), 'failure', and 'retardation' (delay) by placing far more emphasis upon preparation and selection of students for courses, particularly for ONC and OND courses (Peters, 1967). This involved specifying the direct entry requirements mentioned earlier – including 4 O level passes – as well as preparatory courses for those who had not met such requirements. Thus, in 1962, the 3-stage (3-year) structure was replaced with a 2-stage (2-year) structure, preceded by a 1-year or 2-year General course for anyone who did not meet the direct entry requirements. These General (G) courses were intended to be 'diagnostic' and provided a foundation for the careful triaging mentioned earlier (Morrison, 1966).

Unfortunately, reforms instigated by the white paper did not straightforwardly solve problems of wastage, retardation, and failure. Focusing on the situation for Engineering exams, Taylor & Beaumont observed that overall percentage pass rates for City & Guilds exams had hardly changed from 1956 to 1966 (Taylor & Beaumont, 1967). Pearce echoed this concern, noting that pass rates had "flickered upwards by only a few per cent" since the introduction of the revised schemes (Pearce, 1975, page 54).

The practical approach

Even during the early-1950s, it was often said that courses for National awards were "unduly academic" (Foden, 1951, page 43). It was entirely possible for a student to obtain a National Certificate in Engineering, for instance, without ever having been inside an engineering works. Courses for National awards tended to be "stereotyped and narrow" and concerned "mainly with the theory rather than the practice of industrial processes" (Foden, 1951, page 43). Part of the reason for their narrowness was the lack of time available to learners studying part-time (Peters, 1967). However, part of the reason was by design. It was recognised that employees often lacked the underpinning knowledge and understanding that had become increasingly critical to an effective workforce. So, this was what the part-time, day-release courses came to cater for, while practical 'trade knowledge' was generally assumed to be picked up on the job (Clegg & Jones, 1970; Unwin, Fuller, Turbin, & Young, 2004).

If the Crowther committee aspiration for far more technicians to begin their training on full-time college courses was to be realised, then this situation would need to

change. Part of the explanation for their desire to rehabilitate the 'practical' approach would surely also have related to this.

In addition to courses for National awards being unduly academic, the main method of examining for National Certificates and Diplomas continued to be the written exam (Foden, 1966). Foden noted that "the only evidence of practical competence required [...] is the record of laboratory work" (Foden, 1951, page 43).

The situation for craft certificates appears to have been somewhat less biased to the written exam. City & Guilds had grappled with the development of practical exams since the 1890s. Despite often being difficult to administer and invigilate, as well as often being expensive and inconvenient, they had become a significant feature of craft subjects by the 1910s (Foden, 1966). In a section on courses for 'average craft apprentices' Wheatley explained that:

It is regarded as important that success in the examination at the end of the course should have correlation with the student's potentiality as a craftsman; this enhances the importance of proper assessment of practical ability by means of a practical examination or other methods, e.g. course work and specimen work. This practical element also has considerable importance in demonstrating to industry the value of a co-ordinated scheme of industrial training and related further education, especially as no scheme of part-time further education can attempt to provide a substitute for skill training, which is an industrial responsibility.

(Wheatley, 1959, page 42)

The 1964 Industrial Training Act was soon to raise awareness of the importance of effective training, providing a new impetus for integrating the practical approach (Clegg & Jones, 1970).

Continuous assessment

Although it is clear from Foden (1951) that students' homework and classwork marks during their final year were already an important consideration in awarding National certificates, it is also clear from Foden (1966) that innovations during the 1960s embedded these ideas further, including the adoption of new methods of continuous assessment of practical work.

The incorporation of continuous assessment within national certification schemes had been discussed widely for some time. Arguing for reform of the School Certificate exam system, the Norwood report was very positively disposed towards greater reliance upon teacher assessment (Norwood, 1943). Relying on those who knew their students best would enable a much more comprehensive certification process, to paint a broader picture of attainment than was possible with exams. This

would be good for teachers (in terms of professional development) and for students (in terms of fostering a more student-centred approach). Although these ideas did not take root with the introduction of the General Certificate of Education Ordinary level (GCE O level) in the early-1950s, they were a central feature of the Certificate of Secondary Education (CSE) that was introduced during the mid-1960s.

Many assumed that devolution of responsibility for the assessment of student performance was at least as necessary in further education as in secondary schools (for example, Leese, 1966). Although the National award schemes were consistent with this devolutionary approach, Leese claimed that they had ossified, becoming just another externally controlled exam. Whereas, in theory, the REUs might prefer to base results on regionally moderated college assessments, in practice, they ended up paying examiners to set papers and do the marking. Leese argued that this situation ought to be reversed.

Like many others during this period, Leese argued the case for teacher assessment largely in terms of the benefits for teachers and students of following a locally-relevant, teacher-devised syllabus.²⁰ In a subsequent paper, Bacon (1969) also argued the case for increasing reliance on continuous teacher assessment within national certification schemes, albeit with less emphasis on the issue of syllabus control. Instead, he emphasised that continuous assessment ensures that certification is based upon work from across the entire syllabus and not just across the small sample of the syllabus that features in an exam, thereby echoing the Norwood report. Bacon noted that City & Guilds was already committed to extending the contribution of continuous teacher assessment to their Craft Certificates, on the basis of a wealth of experience with practical exams.

Haslegrave report

In May 1967, the Secretary of State for Education and Science invited the National Advisory Council on Education for Industry and Commerce to review the national pattern and organisation of technician courses and exams. The Council appointed a Committee on Technician Courses and Examinations, chaired by Dr H.L.

Haslegrave, which reported 2 years later (Haslegrave, 1969). The committee had a fairly broad remit, which extended beyond technicians to “comparable occupations” (Haslegrave, 1969, page 4). This meant that it covered qualifications designed for commerce as well as for industry, although industry qualifications tended to predominate.

²⁰ The GCE examining boards had experimented with this idea, by providing ‘mode 2’ and ‘mode 3’ syllabus options, which offered increasing amounts of teacher control (compared with the default ‘mode 1’ traditional syllabus option). Despite the strength of the case in favour of increasing teacher control, very few schools actually opted into these schemes (Montgomery, 1965).

The review provided an opportunity to consider the success of reforms arising from the 1961 white paper.²¹ This included changes to the structure of National awards, the introduction of pre-technician General (G) courses, an increase in the number of available subjects for National awards, and a substantial increase in the number of bespoke City & Guilds Technician (T) courses, which had become very popular.²² Alongside National awards, the new T courses offered a less academic route to a technician qualification. G courses assumed a “prognostic” function, routing students one way or the other (Haslegrave, 1969, page 30).

Critical to the context of the review were training reforms that followed the white paper, which had been crystallised in the 1964 Industrial Training Act. These presented the “distinct possibility of profound changes over the whole field of further education” (Haslegrave, 1969, page 20). Indeed, the committee envisaged far greater collaboration between industry (including commerce) and further education, in producing the workforce that the country needed.

Even during the late-1960s, many (if not most) technicians had no relevant qualifications. The committee envisaged a future in which this would be unthinkable, which meant that appropriately specified qualifications would need to be developed. The committee acknowledged the work of the recently established Industrial Training Boards (ITBs) in specifying the kinds of skills that the country needed:

Clearly, a considerable task of job analysis must be undertaken as a first step towards this ideal state of affairs. In our view the ITBs must accept the major responsibility for seeing that this is done. [...] We hope that in doing it, there will be full co-operation with the further education service, both in prescribing suitable courses for technicians of various kinds, and in analysing jobs and devising training programmes.

(Haslegrave, 1969, page 6)

Joint planning, the committee argued, could produce courses that were not only educationally sound, but also reflected the latest training needs. Technical developments in industry – which meant that products were more complex and that production required new applications of science and technology – underscored the

²¹ The white paper addressed 4 main objectives: (i) to broaden students’ education and provide maximum continuity between school and technical college, (ii) to adapt the system more closely to the needs of industry, (iii) to increase the variety of courses available for students, and (iv) to reduce wastage substantially (Haslegrave, 1969). Bourne (1984) argued that the interaction of (ii) and (iii) pointed to the creation of narrowly specialised courses that were tailored to meeting the immediate needs of industry rather than the long-term career prospects of students, and that City & Guilds had responded to this brief by setting up many new technician courses in specialised fields.

²² These new T courses had been developed for industry but not for commerce.

importance of this forward-thinking approach. So, too, did the huge impact that computers were having on clerical and administrative jobs.

Although the committee believed that the new G and T courses had been fairly successful, it was less positive about the reformed National awards. Various adaptations to National courses had improved students' chances of success, including: options for 'lateral' transfer to a course of a different standard without having to completely start again, no longer requiring students to pass exams at the end of each year of their course, and some relaxation of the requirement to pass in all subjects at one sitting (Haslegrave, 1969). Unfortunately, though, drop out and failure rates were still far too high.

The committee raised questions concerning the evolving purposes of these qualifications, it asked whether Ordinary Certificate standards were too high, and it suggested that the Higher Diploma might not even be necessary. Slightly different concerns were raised for the still relatively new Business Studies Nationals. Here, the Ordinary courses were felt to be satisfactory, but they tended to be used as a route to a specialised professional qualification rather than to a Higher course. In fact, the Business Studies OND had proved to be one of the most successful ONDs, appealing to many students, particularly girls, as an alternative to the A level route.

Before drawing conclusions, the committee returned to the continuing problem of drop out and failure (particularly for ONCs) and various potential solutions that had been proposed. These included more time for study, changes in the frequency and type of exams, and the creation of more opportunities for transfer. The committee accepted that students on part-time ONC courses needed more time to complete their studies satisfactorily, albeit also acknowledging that this could be achieved by rationalising the current courses. It also accepted that there should be even more flexible arrangements for transfer across courses, where students had found them to be either too difficult or too easy.

Importantly, the committee noted substantial criticism of the frequency of exams and of some of the methods used. These criticisms emphasised both the "confining influence on the teaching" and the belief that they provided a "false picture of the student and his real achievement" (Haslegrave, 1969, page 42):

Many suggestions were made about different kinds of examination techniques that might be adopted in connection with the examination of technicians. In general, there was a strong trend of opinion towards greater participation by the students' teachers. Most of the following examination components suggested in the evidence give the opportunity for such participation, although the first and most commonly used does not do so:

- (i) a written paper externally set and marked;
- (ii) a written paper set and marked internally;

- (iii) a written paper set and marked internally, with external assessment;
- (iv) an objective test, externally set and marked;
- (v) practical or oral examination, or both, dealt with externally or internally with external assessment;
- (vi) course work assessment by the student's teacher; and
- (vii) project work, internally or externally assessed.

The general view was that technician examinations should include some or all of these components, as appropriate to the case, but that in all cases more weight should be given than at present to assessment by the teacher. In fact, there was a body of opinion that the teacher's assessment should constitute the most important single component in the system of student testing, with other components, externally set or moderated, used as an independent check on the validity of the assessment. This, it was thought, would give the most reliable view of the student's performance and ability.

(Haslegrave, 1969, pages 42 to 43)

The committee noted that students who were better at passing exams were not necessarily those who proved subsequently to be better technicians. Furthermore, technicians needed to be able to solve real-world problems that involved extracting information from multiple sources, analysing it, acting on it, reflecting on actions, and adjusting those actions as necessary. Yet:

The traditional external examination was an unsatisfactory way of testing ability of this kind. What should ultimately be aimed at was an end-of-course "profile" of a student by reference to which both his academic and industrial ability could be gauged.

(Haslegrave, 1969, page 43)

Finally, the committee argued that the weight of evidence from both industry and commerce indicated that the TVET qualification landscape was too complex and insufficiently standardised, being driven by a plethora of controlling bodies that operated without effective co-ordination.

Ultimately, the committee concluded that the reforms set in motion by the 1961 white paper had proved to be insufficient. Moreover, they had facilitated an unco-ordinated proliferation of new courses, which had made the system much harder to understand, let alone to plan, and had frustrated the development of effective educational resources (see also Bell, 1968). In short, more radical reforms were now required.

Summary of the 1960s landscape

Although we argue that the first CASLO qualifications of national prominence were introduced during the late-1980s, the following section will explain how the shift towards outcome-based and mastery-based qualifications actually occurred during the 1970s. So, although the CASLO approach crystallised during the 1980s, and was strongly influenced by the socioeconomic challenges of this period, its roots were grounded in attempts to solve a variety of problems that had become endemic within the TVET landscape by the end of the 1960s. These included:

- the need for more young people to spend more time in post-compulsory education and training (to support the need for skilled craftsmen and technicians)
- serious problems of wastage (drop out), retardation (delay), and failure associated with traditional TVET qualifications
- a desire to rehabilitate the idea of 'practical' education and training as a valued (and motivating) alternative to 'academic' education
- a belief that traditional TVET qualifications (notably Nationals and Higher Nationals) were insufficiently focused on practical skills, and were too dominated by the written exam format
- a belief that greater reliance upon continuous assessment had the potential to improve the comprehensiveness and authenticity of TVET assessment, especially the assessment of higher-level cognitive competencies

1970s

The Haslegrave report laid foundations for radical change in the landscape of TVET qualifications in England. This was led by 2 new bodies, but also by the existing awarding organisations, all of which were exploring new approaches to qualification design. These included outcome-based and mastery-based approaches that foreshadowed the arrival of the CASLO approach (without yet quite constituting it).

Haslegrave recommendations

The principal solution to problems identified in the Haslegrave report was the establishment of administrative machinery capable of co-ordinating and sustaining subsequent reforms on a national basis. Although the committee anticipated that a single national council would be established at some point in the future, it recognised the utility of establishing 2 separate councils in the first instance – one for industry and one for commerce – a Technician Education Council (TEC) and a Business Education Council (BEC). These bodies would be responsible for planning, co-

ordinating, and administering technician and technician-level courses, exams and educational qualifications of a national character:

The TEC would, as soon as possible after its appointment, assume policy and planning responsibility for examinations and qualifications in the whole of the technician field at present covered by joint committees, the CGLI and the REBs. In due course, it would become responsible for syllabuses, assessment, and the award of educational qualifications. It would therefore require a suitable sub-structure of advisory committees to do the detailed syllabus and assessment planning for particular subjects or groups of cognate subjects.

(Haslegrave, 1969, pages 53 to 54)

The committee anticipated a similar role for the BEC, albeit with some difference in emphasis regarding its functions. It recommended that strong consultative and operational links should be established between the 2 councils, and that common assessment policies should be adopted where appropriate. On assessment, the committee stated that:

The award of a technician or comparable qualification, whatever the level, should never depend solely on the student's performance in a formal examination. [...] we would assume that some sort of formal test, externally set or assessed, would continue to feature in all cases. It should not, however, be accorded the same degree of importance as in the past.

(Haslegrave, 1969, page 79)

we think the time has come to move away from the present emphasis on external examinations in favour of the introduction of continuous assessment and other internally applied techniques, with external checks kept to the minimum consistent with the attainment of broad national standards. We wish to see the TEC and BEC give a lead in this direction.

(Haslegrave, 1969, page 61)

Thus, while the councils would have responsibility for national standards, programme design and assessment delivery could largely (or at least partly) be devolved to local providers. In particular, the committee recommended exploring the modular, or credit-based, approach that had recently been developed by City & Guilds:

the Councils might consider in appropriate cases basing the grant of technician and comparable qualifications on studies undertaken under the "credit" system, i.e. the gradual accumulation of passes in subjects which have been studied separately and not as one of several forming a grouped course.

(Haslegrave, 1969, page 79)

Finally, in response to concerns over failure rates and unduly high standards, the committee stated that:

we are strongly of the opinion that any student who fulfils the entry requirements for his particular technician course, and works reasonably hard and well during the course, should be entitled to expect that he will pass the examination.

(Haslegrave, 1969, page 79)

The TEC and the BEC

At the heart of the Haslegrave recommendations lay concern over a lack of national co-ordination, which had led to the proliferation of courses and committees, and had resulted in a “bewildering picture of complexity to employers and students” to the point where even those concerned with providing and administering courses and exams found the system very hard to understand (Haslegrave, 1969, page 44). The TEC and the BEC were established, as independent bodies, to make and direct national policy in their respective fields, to rationalise and simplify provision.

Haslegrave recognised that these organisations would need administrative support, and proposed that City & Guilds should be invited to support them both. This was partly because City & Guilds already had exactly the right kind of machinery in place, dealing with many of the technician candidates who were to fall within the TEC ambit. But it was also so that the kudos of City & Guilds might rub off on the new-style national qualifications. This arrangement might also help to support progression from lower-level City & Guilds craft courses to higher-level technician ones, which had been recognised as a problematic transition for some time.

The TEC was established in March 1973, and the BEC in May 1974 (Birbeck, 1980). Set up as registered companies limited by guarantee, they were originally funded by a grant from the Department of Education and Science (DES), although registration fees would mean that they were able to become self-funding before too long (Roberts, 1988). In 1976, the BEC and the TEC formalised their close relations by constituting a joint committee. City & Guilds initially accepted the invitation to support both organisations, although the BEC severed contractual relations in 1980, and the TEC did likewise in 1981. Stevens stated that the relationship had not been constructive in terms of co-ordinating progression opportunities (Stevens, 1993).

TEC awards

The TEC aimed to replace around 90 City & Guilds or Joint Committee advisory committees with around 22 TEC Programme Committees, to achieve a major rationalisation of the technical qualification landscape (Wheatley, 1976). It developed a suite of awards based upon:

- a TEC Certificate programme (around 12 to 15 units, and 900 hours, typically studied over 3 years by part-time day release) and
- a Higher TEC Certificate programme (around 8 units, and 600 hours, typically studied over 2 years by part-time day release)

This roughly corresponded to the structure and standards of the ONC and HNC that were being replaced.²³ To indicate that learners could study at their own pace, each notional year was defined as a level – with 5 levels spanning the 2 programmes.²⁴ Diploma and Higher Diploma programmes extended the Certificate and Higher Certificate programmes, respectively, providing additional units of similar technical depth. This meant that a Certificate could be converted to a Diploma by completing additional units. Diploma awards comprised roughly twice as many units as Certificate awards.

The idea of a unit was an important feature of TEC awards. Bear in mind that drop out and failure had been huge problems for the outgoing Ordinary and Higher awards. Thus, unitisation was intended to help address this, by replacing an overarching course with a programme of study built from self-contained units, each of which could be passed in their own right on a unit-by-unit basis. The potential of unitisation to help improve completion rates – by flexing to meet the needs of individual learners – was viewed very positively (Like, 1986).

Central to the idea of a TEC award was that it should be designed to meet local needs to satisfy local industries with specific technician jobs in mind. This rolled out as a model in which programmes were intended to be developed locally and validated nationally. Local programme development was supported by national Programme Committees, co-ordinated by broader Sector Committees.

Programme Committees also created 'standard units' (that is, off-the-shelf units) that could be incorporated into a locally developed programme. This was particularly useful for units that were common across programmes, like mathematics and electronics.²⁵ As they were intended to be national awards, the TEC was responsible for setting and calibrating standards, which was achieved by validation and monitoring. Each programme was planned and developed by a college working in partnership with local industry, and then submitted to the TEC for validation. Once validated by the relevant Programme Committee, it could then be delivered and

²³ The information in this section is drawn from a number of publications, including: Pearce (1975), Anson (1978), Blakey & Stagg (1978), Bolton (1978), Birbeck (1980), Riches (1980), Halliday (1981), Bourne (1984), Hunter (1985).

²⁴ TEC awards could be obtained by studying full-time, part-time day, part-time evening, block, sandwich or any combination.

²⁵ TEC relied heavily upon the idea of common units in its attempt to rationalise course provision.

assessed locally. TEC-appointed, regionally-based moderators would visit each college to approve the quality of assessment materials and the standard of student work.

A validated programme might comprise units written exclusively by a local college (or group of colleges) although it might also include standard units. Equally, it might include adapted standard units, or units that were written to include options for students with different progression needs. Units were written with progression in mind, such that, to study a Level 3 unit in a particular topic, a student would be expected to have achieved its Level 2 counterpart. Units were graded pass or merit, but the overall qualification was not graded. Certificates identified individual units by name.

Of particular relevance to the current report, all TEC programmes were specified in terms of learning objectives:

TEC believes that the specification of subject material by specific behavioural objectives gives validating committees information concerning not only the topics to be studied but also the depth to which they are to be studied, and thus gives them more information than the conventional syllabus on which to determine the validity of a proposed programme.

(Bolton, 1978, page 33)

These are written in the form of learning objectives, i.e. they specify exactly what the student should be able to do on completion of the unit, e.g. 'State Ohm's law' or 'Deduce the equivalent resistance of two known resistors connected in parallel'. In other words it is quite clear to the student, the lecturer and the employer what the student is expected to achieve in order to pass the unit. In many respects this changes the emphasis in the role of the teacher from someone helping the student to beat the system-through 'question spotting', etc- to someone working with the student to achieve the specified objectives.

(Riches, 1980, page 365)

An appendix to Hunter (1985) contains the specification for a modified standard unit in Electronics (Level 2) of 60 hours duration. Its content was specified in terms of 6 unit topic areas, each one of which was specified in terms of a small number of 'general objectives' (16 in total across the 6 topic areas). Each of these general objectives was associated with a small set of 'specific objectives'. According to the accompanying guidance, general objectives specified teaching goals, while specific objectives specified the means by which a student should demonstrate their attainment. These objectives were referred to as 'expected learning outcomes'.

The following example illustrates how the second topic area of this unit was specified:

B CATHODE RAY TUBE

7 Knows the principles of operation of a cathode ray tube.

7.1 Labels a diagram of a C.R.T.

7.2 Explains the functions of the following:

(a) electron gun

(b) focus control

(c) intensity control

(d) blanking pulses

7.3 States that deflection can be produced by electric and/or magnetic fields.

7.4 Demonstrates the use of timebases and of vertical and horizontal deflection controls.

(Hunter, 1985, page 285)

Colleges were responsible for developing an assessment plan and for assessing students. They were encouraged to use a variety of methods, including tests within units, end of unit tests, and more extended coursework and assignments. Each assessment would test the set of learning objectives that had been studied in the period since the last one.

As discussed in detail by Halliday (1981), TEC guidance on assessing learning objectives suggested that colleges could adopt one of 2 approaches:

- design the assessment to show mastery of each objective
- design the assessment to show adequate achievement averaged over a block of objectives ²⁶

The guidance noted that most colleges adopted the latter approach. Further guidance suggested that students should be achieving around 50% to pass a unit and around 65% for a merit. In short, although the TEC approach paid more than lip service to the idea of mastery, its stipulations were malleable, to say the least (see also Carter, 2012). As such, we might think of these TEC awards as directly prefiguring the CASLO approach, without quite embodying it.

²⁶ A contemporaneous briefing note on TEC (and BEC) awards expressed the expectation slightly differently: "A student is assessed at regular intervals throughout the unit and is required to pass each assessment" (Bracknell/Wokingham School-Industry Partnership, 1980, page 3).

BEC awards

With a slightly narrower remit, the Business Education Council was established to professionalise the less developed sectors of business and public administration, where the demand for further education and training was less well defined (Field, 2018). As explained by its Chief Officer, John Sellars (1977), BEC engaged a wide variety of stakeholders – further and higher education college staff and students, employers and trade unions, professional bodies, and others – and launched its plans in stages via an initial ‘Consultative Document’ (June 1975), a ‘First Policy Statement’ (June 1976), and ‘Initial Guidelines on the Implementation of Policy’ (May 1977), alongside detailed specifications of core studies for courses leading to BEC General and BEC National awards (October 1977).

The first new BEC awards were introduced in September 1978, with the full suite developed for 16 to 21-year-olds comprising:

- BEC General Certificates – 1 year part-time
- BEC General Diplomas – 1 year full-time or 2 years part-time
- BEC National Certificates – 2 years part-time
- BEC National Diplomas – 2 years full-time or 3 years part-time
- BEC Higher National Certificates – 2 years part-time
- BEC Higher National Diplomas – 2 years full-time or 3 years part-time

According to Deputy Chief Officer, Janet Elliott, the new BEC General awards were designed “primarily as a ‘second chance’, to meet the vocational needs of 16- and 17-year-olds, who did not excel in the school examination system and who have not more than three ‘O’ levels” (Elliott, 1979, page 227). This included students who might previously have attempted a Certificate in Office Studies. BEC National Certificates and Diplomas were phased in from 1978, as the ONC and OND in Business Studies and the ONC in Public Administration were phased out. Similarly, BEC Higher National Certificates and Diplomas were phased in to replace the old HNCs and HNDs. It is worth noting that these replacement HNCs and HNDs retained the same nomenclature, despite being completely new awards.²⁷ It was originally proposed to drop ‘National’ from the title, to avoid confusion, but this policy was reversed to help retain a level of recognition for the new awards (Hannagan, 1978).

Whereas course development for General and National awards was largely centralised – BEC published compulsory core modules for each course and an

²⁷ Morris (1977) estimated that this would have affected about 40,000 candidates per year on HND or HNC and OND or ONC courses in Business Studies, and about 25,000 candidates per year on a variety of other further education courses in the business field.

extensive range of optional modules – Higher awards were essentially designed by colleges and validated by BEC (Elliott, 1979). Having said that, colleges were required to follow BEC guidance on core content and course design. Elliott described as one of the “most interesting features” of the new awards:

the extent to which BEC has required all those involved in teaching and planning business studies courses in the non-degree sector to review their teaching methods and student-learning objectives

(Elliott, 1979, page 227)

More frankly, Morris (1977) characterised BEC policy as the enumeration of a radical change in educational philosophy, which involved:

- manifesting a distinctly vocational purpose for these awards, inviting increased participation from employers, and identifying a prominent role for work experience (new vocationalism)
- structuring all courses in terms of modules, none of which corresponded to traditional subject areas (modularisation) ²⁸
- ensuring that all courses required students to integrate knowledge, skills, and understanding from across a range of disciplines (integration)
- locating 4 themes at the heart of all courses – Money, People, Communication, a Logical and Numerate Approach to Business Problems – to be developed throughout (thematic)
- a move to student-centred, enquiry-based learning (progressivism)

Following a similar path to that trodden by the TEC, the BEC specified module content in terms of learning outcomes. For instance, the following Administration in Business module was issued in 1977, comprising a single general objective and 3 specific ones:

C Understand the importance of the computer as an information tool and be aware of its impact on administrative operations

[This was supported by three learning objectives as follows:]

C1. describe the main characteristics of the computer, including both hardware and software, recognizing the special need for relevant and accurate input data;

C2. identify the main commercial applications of computers from routine data processing to the provision of management information;

²⁸ The modules were intended to occupy from 75 to 90 hours of guided study (Fisher, 1999). Note that BEC ‘modules’ corresponded to TEC ‘units’ (and later became known as units).

C3. outline the way in which specific administrative procedures have changed in response to the introduction of computer systems.

(Fisher, 2003, page 258)

Discussing the development of a Communication module, Pearce suggested that the BEC had adapted its approach to specifying outcomes following criticism of early TEC units. Its objectives were effectively “one step down” in terms of specificity, meaning that the BEC specific objectives were more like TEC general ones (Pearce, 1978, page 7).

BEC policy firmly insisted upon a combination of in-course assessments and terminal exams, although arrangements differed at different levels. For instance, at General level, each core module was assessed by in-course assessment and by an externally set (national) exam paper. Optional modules involved in-course assessment only (Davies, 1981).

At National level, Fisher (2003) characterised assessment arrangements for a National Diploma as follows:

- all general objectives, across all core and optional modules, had to be assessed via in-course assignments
- a student would typically face 9 exams (3 at the end of year 1, 6 at the end of year 2)
- exams were internally set, with approval from an external moderator, and extended case-study exams were encouraged ²⁹
- exams were internally assessed, with external moderation
- there was a strong commitment to criterion referencing – across both in-course assignments and external exams – with a focus on evidencing learning outcomes rather than awarding marks
- modules were graded using a wide range of grades (A to E, or F) but the qualification was graded using only pass and distinction

In-course assignments were fundamental to the new BEC philosophy. They were expected to draw upon abilities developed by the objectives from 2 or more modules

²⁹ Milloy & Saker illustrated the evolution of an end-of-course cross-modular exam, from 1981 to 1983, which became increasingly grounded in real-life problem-solving: “Students were again placed in a fictitious company for two days and were asked to respond to specific situations. These took the form of role plays, memorandums, reports and, for the first time, computer response and group work. People from local business and education were invited to take part throughout the two days. Students brought their set texts, notes and graded in-course assignments and were encouraged to confer during breaks and, if they wished, to re-submit work at any time during the two days.” (Milloy & Saker, 1984, page 24)

of the course (Sellars, 1977), and therefore came to be known as Cross-Modular Assignments. CMAs were specifically designed to help students to integrate knowledge, skills, and understanding from their core studies by applying them to practical business problems. As such, assignments were intended to function both as assessments and as sites of learning.

It is important to recognise how the emphasis on modularisation and integration presented a particular challenge to traditional approaches to teaching for business awards. These had traditionally been delivered on a disciplinary basis, with separate inputs from specialists in economics, mathematics, law, and so on. The new modules incorporated content from different disciplines and the new philosophy invited a quite different approach to teaching and learning that was explicitly premised upon cross-disciplinary integration across modules.

Reception

Both the TEC and the BEC championed radically new approaches. Ellison described the introduction of BEC awards as “a root and branch destruction of the old order” (Ellison, 1987, page 105). Some teachers welcomed these changes. Others bemoaned them. The opinions of scholars also varied. For instance, Franklin, Rawlings, & Craven presented results from a survey of college course leaders, which seemed to indicate that the new awards were failing to achieve their aims and objectives:

we argue that in reality, the old national certificate and diploma courses are being taught in the colleges with a thick BEC veneer applied upon them for external appearances

(Franklin, et al, 1983, page 54)

In response, le Roux (1983) argued that it was unrealistic to expect these aims and objectives to be achieved, in full, so quickly. Conversely, to the extent that ‘real change’ was occurring, even if somewhat slowly, the aims and objectives were genuinely being achieved.

In retrospect, it seems fair to conclude that both the TEC and the BEC seriously underestimated the challenge of bringing teachers up to speed with radically new approaches to curriculum, pedagogy, and assessment, particularly given the expectation that colleges would be responsible for developing, not simply delivering, the new programmes (Morris, 1977; MacRory, Beaumont & Taylor, 1977; Pearce, 1978; Humphreys, 1981; Lysons, 1982; Wilson, 1983; Anderson, 1984; Bourne, 1984; Hunter, 1985). Colleges would inevitably have struggled to catch up, and it should not have been surprising if traditional teaching approaches lingered within the constraints of the new model (O’Sullivan, 1987; Stevens, 1989). The scale of change

must have been daunting, if not overwhelming. Recalling his own experiences of the introduction of BEC awards, Fisher noted that:

Those with a fondness for formal lecturing were appalled and, over the next few years, many would opt for early retirement. Leaving speeches would often include a side swipe at the new courses which had, it would be claimed, lowered academic standards and ushered in a new kind of student who would never have been allowed near college in the “good old days”.

(Fisher, 1999, page 24)

It seems that the TEC approach may have been closer than the BEC approach to prefiguring the CASLO approach, as TEC objectives were specified more tightly than BEC objectives, and the TEC strongly promoted the mastery principle even though it resisted insisting that it had to be applied stringently. The TEC also seems to have been more heavily criticised for embracing an outcome-based approach:

We make a plea for a more complete conception of curriculum development and for the need to find an appropriate role for objectives, where the approach can be regarded as one aid (among others) to design, rather than a strait-jacket on the teacher's perception of what technician education is about.

(MacRory, et al, 1977, page 6)

It is fair to say, however, that similar criticisms were levelled at the new BEC awards:

Objectives like ‘list the main reasons why organisations are formed’ (objective A1) and ‘define the concept of “cost” distinguishing between different types of cost’ (objective E5) encourage an emphasis on description and lower-level cognitive skills [which is] found to encourage rote learning and to provide an inadequate basis for further study.

(Mace, 1980, page 65)

That said, it is clear that the impact of TEC and BEC awards on the further education sector – including the beginning of a shift towards learning outcomes – was profound and long lasting. [Evans \(2009\)](#) described this as “possibly one of the most significant developments” in the sector, providing a massive impetus to staff development.

City & Guilds

In the late 1960s, City & Guilds established an advisory committee that would meet from 1968 to 1969 to advise on a number of issues arising from the Industrial Training Act and the work of the Haslegrave committee (Stevens, 1993). One issue was the conduct of tests of practical competence (see Jones, 1971). Toward the end of 1969, City & Guilds established an Examination Techniques Development Unit,

and a consultancy service in competence testing known as The Skills Testing Service, which developed new approaches to testing industrial skills including the idea of phased testing, which was judged to be particularly important for serving diagnostic and formative purposes.

Longbottom, et al (1973) described a programme of phased testing for trainee craftsmen in the shipbuilding industry, which had been developed in collaboration with City & Guilds. The development process began by using task analysis to identify what was involved in the normal course of production work for each of the main shipbuilding trades. For each identified task, a set of assessment points were then specified, to indicate critical features of effective task performance. This detailed procedural scaffolding helped to ensure that foremen would be able to assume the role of assessors, by observing trainees in action and putting a tick or cross against each of the specified assessment requirements. With the expectation that trainees ought to be able to perform satisfactorily across all of the important features, this was essentially a precursor to the CASLO approach.³⁰

New approaches of this sort were described in detail in a book titled 'Testing Industrial Skills' written by 2 former members of the City & Guilds Skills Testing Service, Alan Jones and Peter Whittaker (1975).³¹ Although most of their examples incorporated a classical approach based upon mark aggregation (in contrast to the CASLO mastery approach) the book emphasised the importance of basing test development upon a clear specification of behavioural objectives. It noted the inadequacy of relying upon 'course content' lists, explaining how they needed to be redescribed, first, in terms of a 'statement of skills' (much like CASLO learning outcomes) and, second, in terms of a 'behavioural specification' (much like CASLO assessment criteria).

During the late 1970s, City & Guilds formulated a new policy on training schemes, based upon:

³⁰ Lacking a clear distinction between learning outcomes and assessment criteria this is probably best considered a precursor to the CASLO approach rather than the approach itself, as we have defined it. Ellis (1979) provided a slightly different example of a similar kind of test developed by City & Guilds, this time assessing the task of grilling steaks. The important features identified in this test were indicated as either desirable or essential. For a candidate to pass, all 15 essential points and 8 out of 12 desirable ones had to be ticked. So, again, this sort of test was an important precursor to the CASLO approach.

³¹ They defined skill as "a complex goal directed sequence of activities with a high level of organization and making extensive use of feedback" (Jones & Whittaker, 1975, page 9), distinguishing between motor, perceptual, and language skills, the latter including basic language skills as well as decision making and planning. They noted that it was "probably generally accepted" (page 2) that written tests of trade knowledge alone were not valid for measuring job competence, and that measures of actual performance were required, whether direct (observations) or indirect (effectively simulations, of higher or lower fidelity).

the move towards a process-competence based approach to technical education as an alternative to the traditional subject-based approach. By September 1980 most of the existing Engineering Craft Studies schemes adopted as much as 12 years previously had been re-stated in terms of learning objectives

(Stevens, 1993, page 143)

City & Guilds continued to implement this policy into the 1980s, and from 1984 to 1985 collaborated with the Chemical Industries Association training organisation (with financial support from the Manpower Services Commission) on the development of new standards of competence.³² Quoting a City & Guilds broadsheet, Raggatt & Williams explained that these standards attested to:

the technical performance expected on the completion of training; the precise criteria by which attainment of performance can be assessed; [and] the conditions under which the performance must be carried out

(Raggatt & Williams, 1999, page 38)

In an article entitled 'Training for Competence' the Development Officer at City & Guilds, Rob Christie, described an emerging zeitgeist:

Fortunately, it is becoming increasingly common practice for the designers of education and training events to specify in clear, behavioural, terms the outcomes which they intend to achieve by the event. Moreover, these outcomes are increasingly likely to be expressed as the results which a worker's behaviour achieves rather than just the activity exhibited. This is an important point for the conception of competence. And particularly if the intended outcomes are skills – cognitive or physical – they are now more likely to be expressed in such a way that they indicate the degree of skill – or the level of competence – expected.

(Christie, 1985, page 30)

Not only did this article emphasise the detailed specification of outcomes, it also stressed the importance of performance testing – either practical or cognitive – and the importance of certifying total mastery of the specified domain. Just a few years later, this mastery-based conception of (training and) certification was to become the foundation for NVQ development. Note, in particular, the reference to how learning outcomes were increasingly being expressed in terms of what the worker's behaviour would achieve rather than just the activity being undertaken. This approach was to become fundamental to the development of National Occupational Standards (Norman Gealy, personal communication).

³² These were standards of competence for workplace assessment, intended to complement (rather than replace) further education qualifications (Norman Gealy, personal communication).

The point of this section is to emphasise that the outcome-based approach to qualification design that was to become the template for building NVQs – which we identify as the first CASLO qualifications of national prominence – was not without precedent. Quite the opposite. The TEC, the BEC, and City & Guilds had all been developing similar approaches during the 1970s. Indeed, the Further Education Unit report ‘Assessment, Quality and Competence’ (FEU, 1986) noted that the BTEC, City & Guilds, and the RSA were all heavily invested in developing outcome-based qualifications during this period, to represent competence more comprehensively and authentically than had been the case in previous decades.

Summary of the 1970s landscape

At this point, it is useful to stand back and survey the Technical and Vocational Education and Training landscape towards the end of the 1970s. Perhaps the most important thing to emphasise is that, while formal qualifications played an important role during the 1970s, they were not as ubiquitous or as important as they are today. The school leaving age had been raised to 16 in 1972, but many young people still entered the job market with few (if any) academic qualifications. Furthermore, many became employed in jobs that provided little (if any) systematic education or training.

Wheatley (1976) emphasised that although all craft apprenticeship college courses led to final exams – notably City & Guilds Craft Certificates – apprentices were generally not required to pass these qualifications to complete their apprenticeship.³³ Traditionally, apprentices merely had to participate in training activities and ‘serve their time’ in order to be considered a craftsman. Other than within a small number of schemes, the overall apprenticeship was not assessed.³⁴ This lack of formal recognition had a negative backwash impact on industrial training, resulting in a situation in which the quality of training was extremely variable, both within and across industries (Wheatley, 1976).

The situation began to change when Industrial Training Boards (ITBs) were introduced to the training landscape. The ITB approach to improving the quality and efficiency of training included identifying training needs and training standards as a basis for designing and validating training programmes. This involved specifying: jobs (title, job description, job title), training programmes (with implications for work-based and college-based provision), and assessment procedures.

³³ Achieving the qualification would certainly have added status, though, and may have led to a salary increase. Indeed, many craft apprentices actually chose to enrol on technician level courses, which emphasises the value attached to off-the-job training and associated qualifications.

³⁴ There were important exceptions to this general rule, which included City & Guilds qualifications for gas fitters, for instance, which certified full competence across both practical and theoretical aspects (Wheatley, 1976).

Nearly all industrial craft occupations had been catered for by 1971, and nearly all involved some form of phased (or staged) testing for diagnostic and formative purposes. To facilitate this, training objectives were “defined in behavioural terms” that specified what the trainee should be capable of, and these specifications led “to corresponding objectivity in the drawing up of phased tests” (both quotations from Wheatley, 1976, page 22). These specifications were developed on the basis of task or skills analysis, which involved deconstructing each craft into a series of component tasks or skills. The phased tests were developed by colleges, by employers, or by awarding organisations such as City & Guilds.

Prior to the ITB schemes being developed, apprentices were unlikely to undergo any systematic programme of on-the-job training. The ITB schemes changed this situation, specifying both training needs and training standards. Inevitably, these new schemes also required awarding organisations to undertake a major programme of syllabus redevelopment for their off-the-job training courses. Wheatley explained that:

In principle, the syllabus content of a course of associated further education is derived mainly from the job specification for the occupation concerned and, more directly, is based on the training and skill specification and the training programme developed from it. [...] this was only rarely possible before the implementation of the Industrial Training Act 1964

(Wheatley, 1976, page 88)

It is important to note how the new ITB schemes incorporated an outcome-based approach to specifying training requirements. As we saw in the previous section, City & Guilds supported this approach, even for theory courses:

Syllabuses in craft theory have normally been set out in traditional ‘content’ form (e.g. ‘Principles of basic woodwork joints’). Increasingly in recent years they are coming to be expressed in behavioural terms (e.g. ‘The student should be able to explain and illustrate the principles involved in the construction of basic joints’). In the case of the relatively new schemes for the building crafts, syllabuses in traditional ‘content’ form are preceded by statements in behavioural terms of the ‘course objectives’; on the other hand, the most recent schemes – for printing crafts – have syllabuses exclusively in the form of ‘course objectives, i.e., in behavioural terms. This is still an area of experimentation and development and there is a good deal of variation in style.

(Wheatley, 1976, page 89)

In short, the roots of the CASLO approach were already deeply embedded in the TVET landscape by the end of the 1970s, led by advances in the training field.

The influence of the ITBs extended beyond craft occupations. For instance, Wheatley noted how most ITBs also published technician training recommendations, providing examples of how firms could prepare their own technician job descriptions, based on task analysis, and then go on to develop suitable training programmes. The ITBs supported higher-level training too. However, whereas the new ITB schemes had begun to certificate the completion of craft apprenticeships – to recognise a satisfactory level of performance across their training programme – this tended not to be extended to higher-level apprenticeships, where the relevant educational qualifications (and training records) assumed greater significance.³⁵

Wheatley ended his review of the state of apprenticeships in England by reflecting upon the many school leavers who did not join apprenticeship schemes. He made particular reference to an influential report on 'Vocational Preparation for Young People' which had been published by the Training Services Agency (of the Manpower Services Commission) in 1975.

The report argued that the current training system was failing in 2 respects. First, there was insufficient investment in training for craft and technician skills. Second, there was inadequate vocational preparation for the 300,000 young people who entered the labour market each year and received little or no training for their work. This included semi-skilled occupations, clerical, commercial, administrative, distributive, and services fields, and a high proportion of occupations that were dominated by women. These were fields where the ITBs had had least impact. Wheatley noted that legislation had provided for further education by day-release for all young people below the age of 18 since 1918. Yet, this had not been implemented. This meant that many young people in employment received little or no systematic on-the-job training, and no systematic off-the-job training.

The landscape within which TVET qualifications were situated during the 1970s can be summarised as follows. First, a substantial amount of technical and vocational education and training occurred through apprenticeships, although numbers had been declining since the late 1960s and this was a cause for concern. The quality and effectiveness of apprenticeship training continued to be highly variable, although the situation had improved through the work of the Industrial Training Boards. This was particularly important for improving the quality and effectiveness of on-the-job training, where the use of task and skills analysis had made training needs and standards far clearer. In the wake of these developments, outcome-based specifications became increasingly popular as the foundation for off-the-job college courses, pioneered by major providers including the TEC, the BEC, and City & Guilds. Off-the-job training courses – delivered primarily by further education

³⁵ Note that, even within these new ITB schemes, craft apprentices did not have to pass their college-based qualification to be awarded the certificate of completion (Wheatley, 1976).

colleges – provided the underpinning knowledge and understanding for apprenticeships. Yet, their importance should not be overstated. They were certainly very valuable in the labour market, but apprentices were generally not required to pass them in order to complete their apprenticeships, particularly within craft industries.

Second, many young people who had left school and entered work had limited or no access to education or training, let alone to qualifications. This theme will be developed toward the end of the next section.

1980s

The previous 2 sections (1960s and 1970s respectively) have explained:

- the landscape of TVET qualifications during the 1960s
- problems that plagued this landscape, and
- how the landscape began to change during the 1970s

The present section provides broader and deeper insights into circumstances surrounding qualification developments in England from the late-1980s onwards. This includes insights into the influence of various North American educational movements, and insights into the sociopolitical context of qualification and assessment policy making. Although the educational movements influenced practices prior to the 1980s – including TEC and BEC initiatives, of course – it was during the 1980s that their influence peaked, as the principle of criterion-referencing became embedded in policies and practices across the board.

Roots in North American scholarship

In the following subsections, we will consider 3 educational movements that originated in the USA but that also became influential in England: the Objectives Movement, the Mastery Movement, and the Criterion-Referenced Measurement Movement. Having explained how these movements influenced adoption of the CASLO approach in England, we will then consider the wider sociopolitical context in England during the 1980s prior to the introduction of NVQs.

It is hard to characterise movements like the following, which have all been influenced by scholars from a variety of backgrounds, working in a variety of contexts, and which have been operationalised in a variety of different ways, including very badly! However, the 3 movements discussed below are interrelated, and the links between them are significant. The following subsections capitalise on

this, highlighting some of the most influential thinkers in each movement, as well as how each of these movements impacted on the next.³⁶

Understanding these movements is critical to answering 3 fundamental questions concerning the genesis of the CASLO approach in England:

1. where did the idea of specifying 'learning outcomes' originate?
2. where did the idea of 'mastering' learning outcomes originate?
3. how did both of these ideas take root in England during the 1970s and 1980s?

Objectives

The roots of the Objectives Movement are often traced back to seminal publications by Franklin Bobbitt (1918; 1924). Yet, the most lucid and straightforward account of the importance of objectives was provided a decade or so later by Ralph Tyler (Stenhouse, 1975). His book entitled 'Basic Principles of Curriculum and Instruction' (Tyler, 1949) has been described as the classic statement of the objectives approach (Kelly, 1982).

Tyler believed that effective instructional planning could not begin until a clear account had been provided of what the instructional process was intended to achieve, in terms of how a student was supposed to change as a result of the instruction. He observed that this critical first step of clarifying purposes (educational objectives) was typically sidestepped.

Tyler

To understand the significance of his contribution to curriculum and instruction, it is important to recognise that Tyler's background lay in assessment, or 'evaluation' as

³⁶ Involving so many movements of such large scale, different accounts will inevitably emphasise different historical pathways. For instance, the general roots of Competence Based Education and Training have been traced in slightly different ways by Davies (1976), Neumann (1979), Brown (1994), and Nodine (2016), to name just a few authors. Links to developments in England during the 1980s have been traced by Tuxford (1989) and more broadly by Burke (1995). There could be no definitive family tree of influences on qualification designers in England during the 1970s and 1980s. However, the influences foregrounded in the present report appear to be particularly salient in making sense of the uptake of the CASLO approach, given the particular shortcomings of extant technical and vocational qualification systems, and given the growing appeal of outcome-based education and training, generally, in England during the 1970s and into the 1980s. For instance, while some might start an account of this sort from Frederick Taylor, the present account starts from Ralph Tyler, particularly given Tyler's influence on work of the Schools Council, in England, during the 1960s (Davies, 1976). Tyler's emphasis on specifying general objectives – which he contrasted with the highly specific objectives favoured by behaviourists – also chimes with the subsequent ambition of NVQ designers to rollout a broad model of competence linked to the Job Competence Model.

he preferred to describe it (Newton & Shaw, 2014). During the early 1930s, his publications focused on the limitations of objective tests in educational contexts. The technology of objective testing had been honed during World War One as a practical tool for allocating recruits to roles in the armed forces. Owing to the simplicity of these tests (including, for example, multiple-choice tests) responses to objective test items could be marked objectively, in contrast to the traditional essay exam, which had been shown to have highly subjective marking. According to Tyler, the problem with applying this format to educational contexts related to the construction of test items, which would typically be derived from a topical outline – a content list – and not from an outline of objectives (Tyler, 1931). Tests constructed on the basis of a content list tended to end up measuring the acquisition of information, but little else:

Often, without recognizing it, test-makers have assumed that all the content treated in a course is to be remembered and that a test of the amount of this material which is remembered by the student is an adequate test of the subject. When the instructors of any college subject formulate their objectives, it is quickly evident that there are other mental processes which students are intended to develop.

(Tyler, 1932a, page 256)

Nowadays, we would refer to this limitation as construct underrepresentation. A ‘construct’ is how we define what our assessment needs to measure, and ‘construct underrepresentation’ indicates that the assessment measures only part of what ought to be measured. What we need is comprehensive, authentic assessment, which is faithful to the entirety of the construct, that is, to all intended learning outcomes. This idea of comprehensive authenticity was central to Tyler’s definition of validity:

the usefulness of the test in measuring the degree to which the pupils have attained the objectives which are the true goals of the subject

(Tyler, 1932b, page 374)

To capture educational objectives comprehensively and authentically, Tyler argued that it was essential to define them, not just in terms of content, but also in terms of behaviour, which Tyler interpreted in a broad sense “to mean any appropriate reactions, physical, mental, emotional, and the like” (Tyler, 1936, page 151). Hence, the idea of behavioural objectives. That Tyler described his approach in terms of behaviour is consistent with his background in assessment, which is concerned with criteria for establishing whether or not educational objectives have been achieved. These criteria always, ultimately, relate to performances – observable behaviours of one sort or another – which might include oral responses, physical demonstrations, written accounts, or suchlike.

Characterising cognition in terms of observable behaviours runs the risk of sounding reductive. In fact, nothing could be further from the truth as far as Tyler was concerned. The whole point of Tyler’s mission was to ensure that high-level objectives – including the least tangible and hardest to describe – were represented as comprehensively and authentically as possible from the outset, as a point of reference for comprehensive and authentic instruction, as well as for comprehensive and authentic assessment:

These educational objectives become the criteria by which materials are selected, content is outlined, instructional procedures are developed and tests and examinations are prepared.

(Tyler, 1949, page 3)

	Nutrition	Digestion	Circulation	Respiration	Reproduction
Understanding of important facts and principles	X	X	X	X	X
Familiarity with dependable sources of information	X				X
Ability to interpret data	X	X	X	X	X
Ability to apply principles	X	X	X	X	X
Ability to study and report results of study	X	X	X	X	X
Broad and mature interests	X	X	X	X	X
Social attitudes	X				X

Figure 1. Use of 2-dimensional chart to represent biological science objectives

Figure 1 adapts part of a table from Tyler (1949, page 50), to demonstrate how objectives can be represented more comprehensively and authentically by identifying

the kind of behavioural change that is anticipated for each element of content identified. In this figure, the rows and columns have been reversed to save space. The content aspects of the objectives (within the subdomain ‘functions of human organisms’) are presented as columns, while the behavioural aspects are presented as rows.

What is clear from using a chart like this is that each area of content can be (and often will be) associated with a wide range of behavioural objectives. Indeed, the process of constructing a chart like this forces its developer to think long and hard about the kind of objectives that really do need to be included (marked by an X) and those that might legitimately be excluded. This becomes the focus for effective curriculum planning, pedagogical planning, and assessment planning.

Classical approach in England

The dangers of construct underrepresentation of the sort identified by Tyler had been recognised in England for as long as exams had been in widespread use (see Latham, 1886, for example). By the 1940s, reform of the School Certificate and Higher School Certificate system was on the cards. The Norwood report, which led to the new General Certificate of Education Ordinary and Advanced level system, recounted concerns such as the following:

The subjects themselves are handled too rigidly; they make little contact with each other or with life or reality or future occupation or interests; examination requirements cast their shadow over all; the acquisition of information is given undue importance; a premium is put on memorisation; power of judgment remains untrained; second-hand opinions pass for knowledge.

(Norwood, 1943, page 10)

Although problems such as these echoed concerns expressed by Tyler in the USA, the Objectives Movement does not appear to have influenced qualification development in England during the 1940s and 1950s. Qualifications continued to be specified only partially, in terms of syllabus content complemented by the exam papers that were released each year. This partial specification of a qualification, in terms of syllabus content and past exam papers, reflects what we refer to as the ‘classical’ approach to qualification design. Rather than specifying educational objectives in terms of both content and behaviours, only content was specified.

Cambridge University Press & Assessment has published a useful archive of [past exam material](#), which illustrates what early O and A level syllabuses and exam papers looked like. During the 1950s, the syllabus for a more technical subject, like O level physics, would simply have listed content. Table 1 reproduces an extract

from the Cambridge [1957 O level physics](#) syllabus, which was 9 pages long, listing 79 items of content plus notes on the scope of each item.³⁷

Syllabus	Notes
1. Measurement of length and of volume.	Both f.p.s. and c.g.s. systems are expected. Candidates will not be asked to <i>describe</i> a vernier or a screw-gauge, but may be expected to use them in the practical examination.
2. Measurement of time by use of the simple pendulum.	A knowledge of the formula relating periodic time to length of the pendulum will not be expected; if required in the practical examination, it will be given.
3. Densities of solids and liquids.	Experimental determination of densities, e.g. by density bottle or by weighing and use of a measuring cylinder, is expected.
4. Pressure in liquids and gases; transmission of fluid pressure; the hydraulic press.	Quantitative formulae required.
5. Boyle's Law.	Experimental demonstration for air is included.

Table 1. Cambridge 1957 O level physics syllabus

By the 1970s, little had changed. The Cambridge [1974 O level physics](#) syllabus was now 12 pages long, but was laid out in essentially the same way. It was more clearly delineated into sections and subsections:

Section A (Items 1 to 20) Mechanics, Hydrostatics, Heat

Section B (Items 21 to 32) Waves, Optics

Section C (Items 33 to 51) Magnetism, Electricity, and Modern Physics

But it was still just a list of 51 items of content. It included a short description of the structure of the exam papers, with a brief introductory section that issued a warning that seemed (ironically) to hint at the perils of not stating objectives clearly:

³⁷ It also provided a description of the practical exam, which would “test whether the candidates have worked through a satisfactory course in the laboratory and are capable of handling simple apparatus” (UCLES, undated, page 37).

The syllabus is not intended to be used as a teaching syllabus, or to suggest a teaching order. It is expected that teachers will wish to develop the subject in their own way.

In the examination, questions will be aimed more at testing the candidates' understanding of fundamental physical principles, and the application of these principles to problem situations, than to their ability to remember a large number of facts and to perform numerical exercises. Some questions will, however, include appropriate calculations.

(UCLES, 1972, page 37)

This lack of detail was characteristic of qualification specification in England during the 1970s and 1980s, including for vocational qualifications (Blakey & Stagg, 1978; Black & Wolf, 1990). This is an important part of the context for the introduction of the CASLO approach, which was intended to help rectify problems associated with the classical approach, most notably its under-specification of educational objectives.

Bloom's Taxonomy

Benjamin Bloom was a student of Tyler. He is most famous for developing and promulgating the behavioural objectives approach, through a book that was to become known as 'Bloom's Taxonomy' (Bloom, Englehart, Furst, Hill, & Krathwohl, 1956).³⁸ The idea for a classification system of this sort had emerged in 1948, as a basis for facilitating communication among examiners. Again, this link to assessment is important: the framework was conceived as a means of facilitating the exchange of test items among university teachers (Krathwohl, 2002).

A core feature of their classification scheme for the cognitive domain was its representation of levels of cognitive complexity, which reflected an assumption that simple behaviours become integrated to form more complex ones. They ordered 6 major classes of behaviour from least to most complex:

1. knowledge
2. comprehension
3. application
4. analysis
5. synthesis
6. evaluation

³⁸ This was published as the first in a series of handbooks, this one focusing on the cognitive domain. Others would focus on the affective domain and the psychomotor domain (although the taxonomy for the psychomotor domain was never published).

The authors of the Taxonomy treated these high-level classes as ‘descriptive’ rather than ‘explanatory’ constructs (Bloom, Hastings, & Madaus, 1971, page 24).³⁹ This oriented them towards further behavioural deconstruction. For instance, they deconstructed the ‘comprehension’ category into subcategories, describing the behaviours associated with each subcategory, thereby helping to render them (and the higher-level category) less covert:

2.00 Comprehension

This represents the lowest level of understanding. It refers to a type of understanding or apprehension such that the individual knows what is being communicated and can make use of the material or idea being communicated without necessarily relating it to other material or seeing its fullest implications.

2.10 Translation

Comprehension as evidenced by the care and accuracy with which the communication is paraphrased or rendered from one language or form of communication to another. [...]

2.20 Interpretation

The explanation or summarization of a communication. [...]

2.30 Extrapolation

The extension of trends or tendencies beyond the given data to determine implications, consequences, corollaries, effects, etc., which are in accordance with the conditions described in the original communication.

(Bloom, et al, 1956, pages 204 to 205)

The intention underlying this process of deconstruction was to unpack the meaning of the higher-level constructs by explaining what they were likely to entail in practice. This was facilitated by the use of “point-at-able” verbs (Bloom, et al, 1971, page 33), such as: to state, to match, to predict, or to compute. Following in the tradition pioneered by Tyler, these verbs explained what students needed to ‘do’ with the syllabus content they were studying. Also following his lead, the most common use of this framework was to secure comprehensive authenticity: to shift curricula and

³⁹ Accordingly, we say that evidence of having solved a particular problem in chemistry permits us to attribute a certain level of understanding to a student (a descriptive analysis of understanding), rather than saying that having a certain level of understanding enables a student to solve a particular problem in chemistry (an explanatory analysis of understanding).

tests away from less complex categories and towards more complex ones (Krathwohl, 2002).⁴⁰

Behaviourism

Before explaining how education scholars in England reacted to the growing influence of the Objectives Movement during the 1970s, it is important to consider how the movement may (or may not) have been influenced by behaviourism. The Pan 'Dictionary of Philosophy' (Flew, 1979) characterises behaviourism as the theory that psychological functioning is definable in terms of observed behavioural data, citing the North American psychologist John B. Watson as its progenitor (Watson, 1925).⁴¹ Although, as a paradigm, its time has now passed, it impacted widely, influencing both philosophy and education in the USA and internationally.

It is important to consider the (alleged) influence of behaviourism because – from the outset and to the present day – critics in England have panned the use of objectives, as though the Objectives Movement was, as a matter of principle, self-evidently misconceived: guilty by association with behaviourism. An early example of this comes from a paper by Bull, who maligned the use of objectives by the TEC and the BEC as a “manifestation of behaviourism” (Bull, 1985, page 74), going on to explain that:

The behaviourist approach which underlies the use of objectives is very suspect. It is based on experiments with animals and the last thing the behaviourists came to study was the actual behaviour of man. Fundamentally, in any case, the behaviourists were not really interested in explaining behaviour, or even learning: they were basically interested in conditioning – and it is debateable whether even animals learn much by conditioning in their normal, natural environment.

(Bull, 1985, page 80)

⁴⁰ The taxonomy was revised nearly half a century after its original publication (Anderson & Krathwohl, et al, 2001). Rather than referring to 'behaviours' (which had often been misconstrued reductively) the new publication referred to 'cognitive processes' and the cognitive complexity dimension was reconfigured slightly: remember, understand, apply, analyze, evaluate, create. (Incidentally, we make no apology for repeatedly referring to 'behavioural objectives' within this section of the present report, as explaining what was originally meant by the term helps to illustrate why those who initially criticised the Objectives Movement for being naively behaviourist were wrong to have done so.) Anderson & Krathwohl, et al, also added another dimension to the revised taxonomy, the 'knowledge' dimension, which transformed it into a 2-dimensional framework. The knowledge dimension ranged from concrete to abstract: factual knowledge, conceptual knowledge, procedural knowledge, and metacognitive knowledge.

⁴¹ Leahey (1992) provides a more detailed and subtle account, which helps to unpack many of the complexities, as well as the disagreements, that underlie this much-mythologised paradigm.

Hyland concurred, claiming that the competence-based approach was “founded on dubious and largely discredited behaviourist principles” (Hyland, 1993, page 66), and stating that:

This specific (and seemingly simple) conception of competence is founded squarely on behaviourist learning principles and suffers from all the weaknesses traditionally identified with such programmes

(Hyland, 1993, page 59)

There are, in fact, 2 important associations between the Objectives Movement and behaviourism, related to measurement and instruction, respectively.

The first association concerns the link between the Objectives Movement and a particular approach to the philosophy of science. We have already seen how behavioural objectives were fundamental to the Objectives Movement.⁴² Yet, this insistence upon specifying objectives in ‘behavioural’ terms is sometimes taken to imply that the Objectives Movement is underpinned by a ‘behaviourist’ philosophy (see Melton, 1997, for instance).

A more accurate explanation is that both behaviourism and the Objectives Movement embraced the concept of operationalism, which had been introduced by the physicist Percy Bridgman (1927).⁴³ This philosophical claim insisted that, when speaking of measuring a particular construct, the construct is synonymous with the operational procedure that is used to measure it (Briggs, 2022). By the end of the 1930s, the principle of operational definition had become a matter of dogma within psychology (Leahey, 1992). What are we measuring when we measure, say, intelligence? Nothing more than the behaviour tapped by the particular intelligence test that we happen to be using.

The high-level ‘behavioural’ objectives described by both Tyler and Bloom were explicitly cognitive – ‘knowledge of’, ‘understanding of’, ‘mature interests’ – which implied that they were not directly observable and had to be inferred. This left them open to conflicting interpretation. By deconstructing and explicating these constructs in more overtly behavioural terms, they became communicable, and therefore assessable. This was a natural route for assessment specialists such as Tyler and Bloom to have pursued, as all assessments bottom-out in performances, or behaviours, of one sort or another. In other words, because they wanted to draw

⁴² The authors of the Taxonomy described behavioural objectives as the “ways in which individuals are to act, think, or feel as the result of participating in some unit of instruction” (Bloom, et al, 1956, page 12).

⁴³ Albeit, in the case of the Objectives Movement, only really during a period spanning the middle of the 20th century.

higher-level inferences concerning competence, they needed to elucidate the lower-level performances that might warrant inferences of that sort.

It was exactly this pragmatic stance that led Bloom and colleagues – in accordance with the zeitgeist of the day – to treat behavioural objectives as though they operationally defined the constructs that they were interested in measuring:

Using operational rather than nominal definitions will make statements of educational objectives clear and easier to communicate to others. Words like “understanding,” “comprehension,” and “appreciation” will take on more precise behavioral meanings and will not be open to various interpretations.

(Bloom, et al, 1971, page 24)

In short, rather than reflecting a deeper commitment to a (now outmoded) philosophy of science, the Objectives Movement adopted the ‘behavioural’ stance for largely pragmatic reasons, to provide a solid foundation for assessment.⁴⁴ The Objectives Movement was never logically bound to the concept of operational definition. But, for a period of time, at least, it did lend it some credibility.

The second, and more significant, association concerns the link between the Objectives Movement and Programmed Instruction, which was an approach that gained traction in the USA during the 1950s and 1960s. The roots of Programmed Instruction can be traced back to World War 2 (and the Korean conflict) when the military looked to North American personnel psychology for inspiration (Bloom, et al, 1971). Soldiers needed to be trained to perform fairly straightforward activities, like assembling and disassembling a rifle, as quickly and efficiently as possible. Psychologists approached this challenge by breaking these macro performances (the activities) down into structured sequences of micro performances, which became the building blocks for instruction. This was known as task analysis.

The renowned behaviourist B.F. Skinner was highly influential in the Programmed Instruction Movement, particularly through his controversial publication ‘Teaching Machines’ (Skinner, 1958). Reflecting on his own daughter’s experience of education, Skinner concluded that schools paid too little attention to principles derived from the scientific study of learning. So, he set out to identify a more effective instructional approach. Programmed Instruction capitalised on the potential of technology to deliver a structured sequence of instructional units, which a learner

⁴⁴ As explained some years later by another author of the Taxonomy, the behavioural approach is the “only viable alternative” when required to assess otherwise unobservable “processes and states” (Furst, 1981, page 442). This is exactly how the Training Agency was later to describe the development of standards for National Vocational Qualifications: “The exercise of developing standards for a particular occupational area is equivalent to developing an operational definition of competence in that area.” (TA, 1988a, page 1).

could progress through at their own rate, involving active engagement with little or no external assistance, providing the learner with immediate feedback concerning the accuracy of their responses (Lockee, Moore, & Burton, 2004). Early incarnations, consistent with Skinner's approach, presented learners with units that were so short, clear, and simple that the probability of error was extremely low, thus facilitating the delicate process of shaping appropriate behaviour (Klausmeier & Goodwin, 1966; Sutherland, 1988).

The link to the Objectives Movement should be fairly obvious. Necessarily, the first step in developing a sequence of programmed instruction involves specification of the intended outcomes of instruction. Robert F. Mager was a leading figure in the movement, and his short, engaging, book on 'Preparing Instructional Objectives' (Mager, 1962) was to become the bible for writing objectives (Lockee, et al, 2004).

So, does this provide evidence that the Objectives Movement was based upon dubious and largely discredited behaviourist principles? No. It simply indicates that a particular behaviourist approach to instruction, Programmed Instruction, was premised upon a precise specification of behavioural objectives. In fact, the kind of objectives required for a behaviourist instructional approach were very "specific ones, very numerous and of the nature of specific habits" (Tyler, 1949, page 42). They bore little resemblance to the generalised objectives preferred by Tyler and Bloom, or to those produced during the 1970s by the TEC and BEC.⁴⁵

It is worth noting that, although Mager's book was originally published under the title 'Preparing Objectives for Programmed Instruction', this was changed within a year to remove reference to Programmed Instruction. Likewise, a preface that originally read "It is assumed that you are interested in preparing materials for auto-instructional presentation" (Mager, 1962, page x) had evolved into "It is assumed that you are interested in preparing effective instruction" by the revised second edition (Mager, 1984, page vi).⁴⁶ After all, the book was about how to write objectives, not about how to implement a particular instructional approach derived from behaviourism.

At the heart of Mager's simple proposal was the idea that objectives should be stated in terms of a desired behaviour, that, is, what the learner will be doing when they

⁴⁵ See Bloom, et al (1971) for an example of the minute level of detail associated with objectives developed for Programmed Instruction, including the structural sequencing of these objectives, which was critical to guiding the instructional process.

⁴⁶ Popham recounts a personal exchange with Mager, in 1961, in which Mager presciently explained that once "all the furor about teaching machines and programmed instruction had died down, the single most important contribution of the movement would be the attention it directed to the form in which objectives should be formulated." (Popham, 1978, page 14).

demonstrate their learning. According to the revised second edition, this provided answers to 3 questions (see Mager, 1984, page 87):⁴⁷

1. What do I want students to be able to do?
2. What are the important conditions or constraints under which I want them to perform?
3. How well must students perform for me to be satisfied?

This is an example of an appropriately written objective according to Mager:

Without regard to subject matter or grade level, be able to describe ten examples of school practices that promote learning and ten examples of school practices that retard or interfere with learning.

(Mager, 1984, page 108)

In this example, the performance element concerned 'describing', the conditions involved 'any subject and any grade level', and the criterion for both categories was 'ten examples'. Although disassociated from the conceptual baggage of Programmed Instruction, it is clear that objectives of this sort were still far more specific than Tyler's generalised ones.

In conclusion, although there are links between the Objectives Movement and behaviourism, they have certainly been overstated and overgeneralised. The idea that the movement is somehow fundamentally undermined by association with behaviourism is misguided.

Academic debate in England

When the TEC, the BEC, and (later) the National Council for Vocational Qualifications (NCVQ) began to apply principles from the Objectives Movement to the specification of VTQs in England, they immediately became the target of heavy criticism from scholars of education. MacRory criticised early TEC innovations, warning of the "incomprehensible" "mystique" of objectives (MacRory, et al, 1977, page 4). Bull claimed that objectives of the sort adopted by the TEC and the BEC were "inimical to the real structure of knowledge" (Bull, 1985, page 77). Norris, questioning the new emphasis upon competence – in the wake of the De Ville report, which led directly to the development of NVQs – argued that competence models "distort and understate the very things they are trying to represent" (Norris, 1991, page 334).

⁴⁷ This was originally stated as: (a) Identify and name the over-all behavior act. (b) Define the important conditions under which the behavior is to occur (given and/or restrictions and limitations). (c) Define the criterion of acceptable performance. (See Mager, 1962, page 53.)

Significantly, all 3 of these early critiques referenced Lawrence Stenhouse. The point, here, is that – even before the TEC, the BEC, and the NCVQ began to apply principles from the Objectives Movement to qualification design – the movement had already received substantial criticism from education scholars in England, who seemed intent on heading the Objectives Movement off at the pass in its march from the USA to England. The book ‘An Introduction to Curriculum Research and Development’ (Stenhouse, 1975) was particularly influential in this respect.

Kelly (1982) provides an informative overview of this period, explaining that there had been little interest, in England, in specifying objectives until the mid-1960s. Yet, as problems of curriculum planning came to the fore, interest began to grow, particularly under the aegis of the Schools Council, which was established in 1964. Most of its projects began with the development of clear course objectives.

Although the objectives approach had its supporters in England, including Hirst (1969), other scholars were more critical. They included Pring (1971), Sockett (1971), and Ormell (1974), all of whom criticised Bloom’s Taxonomy specifically.⁴⁸ Wesson (1983a; 1983b) later criticised the use of behavioural objectives more generally, referencing TEC developments specifically. Kelly (1982) concluded that criticisms of the objectives model were “as strong as, if not stronger than” the case for its use (Kelly, 1982, page 108). In fact, to many scholars of education, the case against the objectives model seemed incontrovertible.⁴⁹

An article by Christopher Ormell provides an interesting perspective on this period, written by one of the original critics, albeit some decades later (Ormell, 1992). He described how the Stenhouse critique became the “official story” among progressive academic educationists, such that opposing behavioural objectives became, after 1975, the “badge of progressive educationalism” the world over (Ormell, 1992, page 23). By the 1990s, however, the tide had turned. The progressive principles of the 1970s – open problem solving, value free approaches, creativity, optionality, child-centred work, culturally permissive approaches – now seemed out of date. What was needed was a new case against Bloom, as the old case, the Stenhouse case, was totally ineffective now.

Ormell directly challenged both of Stenhouse’s principal objections to behavioural objectives. First, Stenhouse claimed that objectives provided a straightjacketed account of knowledge, as though knowledge could only be demonstrated in a discrete, pre-specified, manner. Ormell replied that this critique overstated the

⁴⁸ See Furst (1981) for an illuminating response from one of the team that originally produced Bloom’s Taxonomy.

⁴⁹ Having said that, even within these circles, it was often accepted that outcome-based approaches can work very well in courses that focus on training rather than education, that is: “in courses which are essentially vocational” (Kelly, 2009, p.86).

significance of creative (unpredictable) performances. Second, Stenhouse claimed that objectives provide a straightjacketed account of education, as though we should be able to specify in advance, with some clarity, what students ought to learn. Ormell replied that, nowadays, it seemed inconceivable that we should not even try to clarify educational objectives. In short, according to Ormell, the Stenhouse critique embodied values “possibly accepted in the 1970s, but certainly out of favour now” such that there is “no mandate today for unpredictable students and obscurantist teachers” (Ormell, 1992, page 27).⁵⁰

The purpose of this section is not to try to do justice to the arguments for and against the Objectives Movement, whether prior to the 1970s, during it, or subsequent to it. The purpose is simply to illustrate the nature of academic debate among academic educationists in England during the 1970s and 1980s regarding the Objectives Movement. This provides an important backdrop to the adoption of the CASLO approach by assessment organisations from the 1970s to the 1990s, and to how the NVQ model, in particular, was received by many scholars of education.

Mastery

The Mastery Movement became influential in the USA during the 1970s, as both a philosophy of, and a methodology for, teaching and learning. Two parallels with the history of the Objectives Movement are worth noting. First, the roots of mastery learning can be traced back to influential North American scholars working during the early decades of the 20th century. Second, mastery learning was adopted as an organising principle of the Programmed Instruction Movement. The Mastery Movement itself, however, began during the 1970s, in the wake of a report written by Benjamin Bloom, entitled ‘Learning for Mastery’ (Bloom, 1968). According to Gagne – who had previously been associated with the Programmed Instruction Movement and who later developed his own version of mastery learning – Bloom raised the idea of mastery “to a new level of generality” (Gagne, 1988, page 108).

Bloom

The idea of mastery represented a new philosophy of teaching and learning because it rejected the standard assumption that: for each new cohort of students, only about a third will adequately learn what has been taught, about a third will learn a good

⁵⁰ Ormell’s alternative case against Bloom (his approach to specifying behavioural objectives) argued for “whole” objectives for education that “encompass both behavioural and mental objectives” (page 30). He argued that we want students who “actually do understand, actually do think, actually do take safety utterly seriously in the laboratory” (page 31). These students “do the appropriate things, as well as possess the mental states” (page 31). That this should be offered as an alternative to Bloom seems a little odd, to say the least.

deal but not enough to be considered good, and a third will fail or just get by (Bloom, et al, 1971).⁵¹ This assumption was embodied in the standard practice of grading on the normal curve, which led to the highest achieving students receiving the highest grades, and to the lowest achieving students being failed. This was not so much a problem with assessment – the lowest achieving students might legitimately have been categorised as having failed. Instead, what was at fault was the expectation of failure, which created a self-fulfilling prophecy of failure. Bloom argued that this assumption was not simply wasteful and destructive but unnecessary. Conversely, he claimed that:

Most students (perhaps more than 90 per cent) can master what we have to teach them, and it is the task of instruction to find the means which will enable them to master the subject under consideration.

(Bloom, et al, 1971, page 43)

Drawing on work by Carroll (1963), Bloom argued that this was possible as long as the quality of instruction was high enough, and as long as students who needed additional time were provided with it. Indeed, he proposed a relationship between these 2 variables: with effective instruction, we can reduce the amount of time required by slower students to the point where this is not prohibitively long.

Without wanting to be too prescriptive, methodologically, Bloom recommended the approach that he and colleagues had been developing at the University of Chicago. Central to this approach was the idea of formative evaluation.⁵² Starting from a clear and comprehensive specification of learning outcomes, a course could be broken down into units of learning of perhaps a week or two in duration. These units could then be broken down into a number of elements, and diagnostic progress tests could be developed to determine whether a student had mastered the elements, or if not, then what they still needed to learn.

Frequent formative evaluation, the Chicago group argued, helped to pace student learning, and helped to motivate students. For students who had mastered a tested element, formative evaluation would help to reinforce their learning. For the remaining students, the test would provide critical feedback to reveal their particular points of difficulty. Upon the foundation of this diagnosis, a teacher would then prescribe an appropriate instructional intervention to help close the gap in learning.

One obvious challenge associated with this personalised approach to teaching and learning – which was premised on the idea that pace of progression will differ across

⁵¹ Bloom's 1968 report was reproduced, with minor editorial amendments, in both Block (1971) and Bloom, et al (1971).

⁵² The idea of formative assessment, which is now internationally recognised, can be traced back to their 'Handbook on Formative and Summative Evaluation of Student Learning' (Bloom, et al, 1971).

students – is how it can be accommodated when teaching whole classes. Although this might be achieved in various ways, Bloom recommended that enrichment, or extension, activities should be used with faster students, enabling them to deepen their learning while slower students were still acquiring the required breadth of learning (Guskey, 2023).

Central to the idea of mastery learning is the impact that it is presumed to have upon the slowest learners within any cohort. This impact derives from using evaluation formatively, that is, integrating assessment within teaching and learning rather than concentrating it all at the end of a course. If students are supported to achieve mastery in this fashion, then summative evaluation (summative assessment) should become a positive, reinforcing experience:

If the system of formative evaluation (diagnostic-progress tests) and summative evaluation (achievement examinations) informs the student of his mastery of the subject, he will come to believe in his own competence.

(Bloom, et al, 1971, page 56)

Classical approach in England

England and the USA have always had quite different assessment cultures. For instance, England never bought into multiple-choice testing with quite the same fervour as the USA. Yet, below the surface, their working assumptions and models have actually remained quite similar in many ways, and this was certainly true of grading practices during the middle of the 20th century. As such, it should not be surprising that the 2 nations experienced similar assessment and learning challenges, including how best to recognise success and to prevent failure.

Pedley, for example, described arrangements for grading regional technical exams in England, during the 1960s, as follows:

In all subjects the pass mark is 40 per cent [...] In most subjects the credit mark is 65 per cent and the distinction mark 85 per cent.

(Pedley, 1964, page 154)

So, the idea of generally mastering the domain of learning was not built into the grading model for these TVET qualifications, as might have been indicated by specifying a high pass mark. Moreover, we have already seen how failure was a major concern for TVET qualifications during the 1950s and 1960s. Taylor & Beaumont noted that the typical failure rate for a City & Guilds or Regional Examining Board exam during the 1950s and 1960s was approximately a third (Taylor & Beaumont, 1967). Very similarly, O level pass rates tended to fluctuate around the 60% mark, from the 1950s to the 1970s, while A level pass rates tended to fluctuate around the 70% mark (Newton, 2022).

Neither were these exams designed to certify mastery of specific elements of competence. Instead, they were designed according to the classical approach, whereby exam marks were aggregated to a mark total, with candidates' final grades determined by how many marks they achieved in total. That is, these exams adopted a compensatory (as opposed to a mastery) approach to aggregation.

In short, all of the challenges that led Bloom to propose a new philosophy of teaching and learning were just as evident in England during the 1970s as they had been in the USA.

Criterion-Referenced Measurement

In the introduction to his 1978 book 'Criterion-Referenced Measurement', Ronald Berk described the shift from norm-referenced to criterion-referenced testing as the "most dramatic" to have occurred over the past decade in the field of educational measurement and evaluation in the USA. He explained that an "increasing emphasis on mastery-proficiency-competency is permeating all levels of education and other professions, particularly medicine and the allied health fields" (both quotations from Berk, 1978, page 3).

The distinction between norm-referenced and criterion-referenced measurement had been introduced during the early 1960s, by Robert Glaser. Glaser had been a leading figure in the Programmed Instruction Movement and – like both Tyler and Bloom – his expertise in teaching and learning was informed by his background in educational measurement.

Glaser observed that most existing educational attainment measures were norm-referenced, that is, they embodied relative standards, indicating the proficiency of any particular student relative to their peer group. He argued, instead, for criterion-referenced measures, which embodied absolute standards, to indicate the proficiency of any particular student along a "continuum of knowledge acquisition" (Glaser, 1963, page 519). In addition, he argued, that educationists needed to:

specify minimum levels of performance that describe the least amount of end-of-course competence the student is expected to attain, or that he needs in order to go on to the next course in a sequence.

(Glaser, 1963, page 520)

The link to the Mastery Movement was quite explicit. As Jim Popham put it some years later, when the intention is to bring large numbers of learners to levels of competence not previously seen, relative comparisons are no longer meaningful because we want all learners to end up performing at a high level (Popham, 1994). Glaser's take on summative assessment was very similar to Bloom's take on formative assessment. Both involved new ways of thinking about assessment, driven

by new ways of thinking about teaching and learning, plus the need for clarity concerning what students have actually learnt, or not yet learnt.

Berk (1978) explained how Glaser's original conception of criterion-referenced measurement had subsequently been developed (in the USA) in 2 different directions: domain-referenced measurement and mastery testing. Mastery testing was closely linked to the idea of mastery learning, whereby tests were designed to measure particular instructional objectives, and cut-scores were established to distinguish between learners who had mastered those objectives versus those who had not yet mastered them. This had been the dominant direction of travel.

Domain-referenced measurement described the other (more complicated) direction of travel. It aimed to take testing beyond traditional educational objectives, to provide what Jim Popham described as an "unambiguous" definition of each domain of learning (Popham, 1978, page 13). Developing this line of reasoning, Popham had concluded that traditional objectives were simply too vague. He argued that the first step in developing any criterion-referenced test ought to involve a more precise definition of the domain of content or behaviours that needed to be assessed, which could be defined operationally as the specification of all possible test items.⁵³

In fact, neither of these 2 directions of travel strongly influenced developments in England. What was influential, however, was the more general idea of moving away from norm-referencing and towards criterion-referencing. We have already seen how this influenced the work of the technical and vocational awarding organisations during the 1970s, including the TEC, the BEC, City & Guilds, and the RSA. But the idea of criterion-referencing became increasingly popular during the 1980s, with scholars and politicians alike, and came to influence qualification and assessment practices in England far more widely. Ultimately, the ideas that drove the development of technical and vocational qualifications in a certain direction during the 1970s – the direction that was later to take root in the CASLO approach – ended up driving the development of general qualifications in a somewhat different direction. The important point, however, is that practices evolved across the board under the influence of criterion-referencing during the 1970s and 1980s. In the following subsections, we will see how the legacy of criterion-referencing extended well beyond the TVET landscape.

Records of Achievement

A speech by the Secretary of State for Education and Science, Sir Keith Joseph, encapsulated the zeitgeist of the early 1980s. Joseph announced a variety of new policy goals that were premised on more clearly specified educational objectives and

⁵³ Determining an appropriate level of precision proved to be the most challenging aspect of this approach (Popham, 1978; Popham, 1994).


a move towards criterion-referenced assessment. This included the introduction of Records of Achievement:

But despite these difficulties no one of us can be satisfied with what our pupils attain by the time that they are allowed to leave school. Some of these attainments are not at present systematically assessed or acknowledged where they can be and ought to be. That is why I see as important the development of records of achievement on which I have recently issued a draft statement of policy.

(Joseph, 1984, page 139)

This draft statement inspired a conference in 1984, which resulted in a book that was edited by Patricia Broadfoot (1986). In her introduction to this book, Broadfoot explained that there was now a considerable groundswell of support among educationists for “a more comprehensive and curriculum-integrated approach to assessment” (Broadfoot, 1986, page 2), an idea that Records of Achievement clearly embodied. Indeed, she characterised England as being on the brink of an assessment revolution, confronting a longstanding tradition of overreliance upon external exams. Having said that, Broadfoot acknowledged that, while trailblazer schemes from the 1970s and early-1980s had now become a matter of government policy, there were still deep divisions in the movement concerning how best to achieve their mutually accepted goals.

One of these deep divisions concerned the degree to which Records of Achievement ought to be subjective and personal versus objective and comparable. Chapters within the edited book illustrated both extremes. Located at the more objective end of this continuum were schemes that City & Guilds had developed in the wake of the Further Education Unit report ‘A Basis for Choice’ (FEU, 1979), which were firmly grounded in criterion-referencing. Nick Stratton, a Senior Research Officer at City & Guilds, described the schemes that they had developed with particular reference to a general vocational preparation course known as ‘course 365’ (Stratton, 1986).



Practical & Numerical	Safety	Can explain the need for safety rules	Can remember safety instructions	Can spot safety hazards	Can apply safe working practices independently	Can maintain, and suggest improvements to, safety measures
	Using equipment	Can use equipment safely to perform simple tasks under guidance	Can use equipment safely to perform a sequence of tasks after demonstration	Can select and use suitable equipment and materials for the job, without help	Can set up and use equipment to produce work to standard	Can identify and remedy common faults in equipment
	Numeracy	Can count objects	Can solve problems by adding and subtracting	Can solve problem by multiplying and dividing	Can calculate ratios, percentages and proportions	Can use algebraic formulae

Figure 2. Extract of profile grid from Stratton (1986, pages 110 to 111)

City & Guilds was keen to provide a scheme that would record student progress in a manner that would support and motivate students, as well as culminating in a reliable end-of-course attainment profile. Course 365 embodied this idea of capturing both progress and end-of-course attainment by using a profile grid. The (mark 3 version of this) grid was split into 4 generic outcome areas: communication, practical & numerical, social, and decision-making. Rows within each outcome area identified more specific objectives, such as, for the 'social' outcome area: working in a group, accepting responsibility, and working with clients. Alongside each of these rows were 5 columns that exemplified progress in the form of criterion statements that increased, from left to right, in terms of autonomy, complexity, and variety of application (albeit with the caveat that they were not necessarily organised in a strict logical hierarchy). Figure 2 recreates an extract from this profile grid relating to the 'practical & numerical' outcome area.⁵⁴

Clearly, this was an outcome-based approach to recording progress and end-of-course attainment, and each of the criterion statements within this profile grid was amplified using concrete examples. City & Guilds saw each of the 5 criterion statements associated with each row as a 'stepping stone' towards maturity, rather than a formal level or grade.

Although perhaps not quite a direct precursor to the CASLO approach, the City & Guilds profiling schemes clearly reflected a similar ancestry, and were designed in the same spirit. It is worth noting that the Certificate of Pre-Vocational Education (CPVE) qualification model was based upon essentially the same kind of profiling scheme:

It embraces a fully fledged formative and summative profiling system based on both a personal reviewing system and a bank of summative 'can do' statements. The structure of this bank reflects the ten core areas of CPVE. There will be a dozen or so statements for each area and these will be organised as several sets of hierarchically related statements, with some left over 'stand alone' statements. Thus each printed-out profile report will contain only those statements corresponding to best performance.

(Stratton, 1986, page 124)

⁵⁴ The profile grid actually incorporated an initial 'half' column, which acknowledged that some students (often those with a learning difficulty) would finish the course still working towards the first criterion statement for one or more of the objectives.

Grade-related Criteria

In the same speech, Joseph emphasised that external exams – including the soon to be launched General Certificate of Secondary Education (GCSE) – should also move towards a greater degree of criterion-referencing:

The existing system tells us a great deal about relative standards between different candidates. It tells us much less about absolute standards. [...] We need a reasonable assurance that pupils obtaining a particular grade will know certain things and possess certain skills or have achieved a certain competence.

(Joseph, 1984, page 142)

His broader policy goal was to raise educational standards, and there were concerns that norm-referenced grading made it impossible to determine whether or not educational standards were actually rising or falling, whether at the national level or for individual schools.⁵⁵ He believed that a move towards criterion-referencing would render results capable of measuring improvements in educational standards over time. Results would also be better designed to hold individual schools to account, should they fail to raise educational standards.

In fact, his 1984 announcement largely reiterated sentiments expressed in a DES policy statement from 1982. This statement had acknowledged attempts to develop GCSE Grade Descriptions, which were intended to indicate the likely levels of competence and the knowledge that might be expected from those who obtained a particular GCSE grade in each subject area. The Secretary of State hoped that these Grade Descriptions would be a step towards the longer-term goal of developing Grade-related Criteria that would render: “the award of all grades conditional on evidence of attainment in specific aspects of a subject” (DES & WO, 1982, page 10). The exam boards had been asked to direct their attention to this longer-term goal.

Despite a protracted period of research and analysis, the longer-term goal of establishing Grade-related Criteria was never achieved. It was concluded that strong forms of criterion-referencing were incompatible with public examining in England (Cresswell 1987; Gipps, 1990; Tattersall 2007). The boards were, however, successful in developing Grade Descriptions, which helped to exemplify attainment standards at the subject level for both GCSE and A level exams (Kingdon & Stobart, 1988; Kingdon, 1991).

⁵⁵ In fact, O and A level exams had never been norm-referenced, despite what many stakeholders had presumed (see Newton, 2022, for a more nuanced analysis)

Assessment Objectives

Another legacy of enthusiasm for criterion-referencing during the 1980s was the development of assessment objectives, although this can equally be seen as a longer-term legacy of enthusiasm for educational objectives during the 1970s. The Schools Council appears to have been particularly influential in this respect.

The Schools Council was established in 1964 to assume responsibility for most of the work previously carried out by the Secondary School Examinations Council and the Curriculum Study Group of the Department of Education and Science (see Schools Council, 1965). This included a co-ordinating and advisory function in relation to O and A level examining. It very soon decided that rapid changes in schools and society demanded “a complete reappraisal of the sixth-form curriculum and examinations” (Schools Council, 1972, page 7). A paper on sixth-form examining methods indicated a need to revisit the fundamental principles of examining:

The traditional pattern of examinations based primarily on syllabus content has plainly undesirable and constrictive effects on teaching and learning. A valid test of the success of pupils in following a course of study requires not merely that it should test content but that above all it should be related to the aims and emphasis of the teaching that preceded it. [...] Certainly the analysis of educational aims would seem to be the prerequisite of examination reform.

(Schools Council, 1968, page 6)

The concern, here, was that existing exams focused too much on testing “factual knowledge” and too little on testing the “ability to think” (Schools Council, 1968, page 6), with an inevitable negative backwash impact on teaching and learning. In response to this report, the Schools Council invited all of its subject committees to reconsider their examining techniques (Schools Council, 1973).

The report from the science committee noted that curriculum changes had progressed more rapidly in the sciences, influenced by the Nuffield Science projects (see Schools Council, 1970, for additional insights). It concluded that it was no longer acceptable for teachers and examiners to rely purely on syllabus content lists and past paper precedents. Its formal recommendations on the incorporation of clear and detailed objectives included:

The objectives should be explicit and should match the objectives of the curriculum.

The move towards higher mental objectives and away from questions demanding only the ability to remember should continue.

(Schools Council, 1973, page 44)

The science committee welcomed the introduction of ‘examination specifications’ that were based upon the idea of a 2-dimensional chart from Tyler (illustrated earlier in Figure 1). Appendix E from their report included a specification that had been developed for the 1969 Nuffield A level chemistry exam – which represented a full range of topic areas (as columns) alongside a full range of Bloomian cognitive behaviours (as rows) – with cell values indicating intended weighting. Already, these specifications were shifting emphasis “away from the ability simply to recall towards the ability to comprehend and apply” (Schools Council, 1973, page 25).

An authoritative review of O and A level physics syllabuses toward the end of the 1970s underlined the extent to which they differed in style across the 8 boards (Crellin, Orton, & Tawney, 1979). This report also noted that, while the examining boards had certainly engaged with the debate over objectives, they tended to prefer less detailed specifications: rather than developing long lists of behavioural objectives, they preferred more holistic approaches. For instance, a publication by the Joint Matriculation Board (JMB, 1970) included an illustration of how the objectives of a science exam might be represented in terms of 6 dimensions and associated weightings:

Knowledge	40%
Comprehension	30%
Application	20%
Evaluation and investigation	10%
Expression	
Experimental skills	

In this example, expression was not weighted independently as it would be taken into account across all of the questions, and experimental skills would be dealt with separately in the practical exam. This approach is quite similar to how assessment objectives are expressed nowadays, although it took some time before this became standard practice.

A step in that direction occurred during the early 1980s, with the specification of Common Cores for A level subject areas (GCE Boards, 1983). These specifications indicated what all subject syllabuses ought to have in common, restricted to not more than half of any particular syllabus (Kingdon, 1991). Common Cores were intended to be useful for higher education selectors – clarifying what might be expected of prospective applicants – as well as having the potential to improve comparability of standards within and across examining boards. This was in the wake of widespread concern over the proliferation and increasing divergence of A level syllabuses.

Subject working groups were given considerable leeway in developing their Common Cores and – of relevance to the present report – some of them explicated aims and

objectives as well as content. For example, as the first of 3 aims, the geology working group specified that the core syllabus will: “provide knowledge and practical experience of geology, both in the field and the laboratory, and demonstrate the use of this knowledge and experience”. In a subsequent section on objectives, it stated that students will be expected to demonstrate the ability to “perform basic tests and use elementary techniques in the field and the laboratory” and to “describe and understand geological processes both present and past” and to “formulate and test hypotheses in a geological context” and so on (all quotations from GCE Boards, 1983, page 72). Over time, A level syllabuses incorporated increasingly clear statements of aims and objectives (Kingdon, 1991).⁵⁶

Specifications of this sort were developed more systematically within similar regulations for GCSEs, the National Criteria (DES, 1985). The Criteria specified a common structure for the major GCSE subject areas: permitted titles, general aims, assessment objectives, proportions of marks allocated to those objectives, schemes of assessment, and descriptions of standards at key grades (Kingdon & Stobart, 1988). In accordance with their roots in the Objectives Movement, GCSE assessment objectives were intended to indicate that GCSEs were about more than mere recall of factual knowledge, emphasising the importance of higher-order abilities and skills (Butterfield, 1996).⁵⁷ Indeed, part of the rationale for putting coursework at the heart of the GCSE model was to improve the assessment (and thereby the teaching) of these higher-order competencies. Significantly, though, GCSE assessment objectives were intended to be used for designing syllabuses and assessment procedures, rather than to be used directly by teachers when assessing students (Butterfield, 1996).

As far as GCSE and A level exams were concerned, assessment objectives had a significant role to play in warranting comparability claims across the examining

⁵⁶ Despite substantial progress in developing aims and objectives, it is interesting to note how the 1988 Higginson report on A levels echoed exactly the same concerns as had been expressed 2 decades earlier by the Schools Council: “As we have said, there is a need for leaner syllabuses in which the proportion of factual content has been reduced and in which the accent is on higher level skills and making sense of the facts” (Higginson, 1988, para. 5.2).

⁵⁷ The following quotation illustrates how GCSE assessment objectives were originally formulated (reproducing a quotation from the National Criteria for English): “The Assessment Objectives in a syllabus with the title English must provide opportunities for candidates to demonstrate their ability to: (i) understand and convey information; (ii) understand, order and present facts, ideas and opinions; (iii) evaluate information in reading material and in other media, and select what is relevant to specific purposes; (iv) articulate experience and express what is felt and what is imagined; (v) recognise implicit meaning and attitudes; (vi) show a sense of audience and an awareness of style in both formal and informal situations; (vii) exercise control of appropriate grammatical structures, conventions of paragraphing, sentence structure, punctuation and spelling in their writing; (viii) communicate effectively and appropriately in spoken English.” (Abbott, McLone, & Patrick, 1989, pages 3 to 4).

boards. Scrutinising syllabuses through the lens of both breadth and depth – supported by tools like Bloom’s Taxonomy – had raised significant comparability concerns. For instance, an analysis by Crellin, et al (1979) suggested that O level physics syllabuses were fairly similar in terms of breadth of coverage, but varied considerably in the depth with which particular topics were treated. Based on a similar analysis for A level physics, they found it difficult not to conclude that the syllabuses differed significantly in demand. If, instead, all of the boards were required to allocate a certain proportion of marks for knowledge, comprehension, application, and so on, then this would help to deflect criticisms of this sort. Thus, GCSE and A level assessment objectives came to function as a tool for calibrating standards, rather than as a framework for teaching, learning, and assessment.

National Curriculum Assessment

Plans to assess the new national curriculum can be understood as the peak of enthusiasm for criterion-referencing in England. In December 1987, Professor Paul Black submitted proposals from the Task Group on Assessment and Testing (TGAT) to the Rt Hon Kenneth Baker MP. The TGAT report began by explaining that certain design principles would be prioritised: assessments should be criterion-referenced, they should be used formatively, and they should relate to progression through the curriculum:

More generally, the combination of a norm-referenced system with age-specific scaling would not be consistent with the proposals in the national curriculum consultative document. The overall national purpose is to work for achievement of the attainment targets of the curriculum. Assessment, whether for feedback to pupils or overall reporting and monitoring, should therefore be related to this attainment i.e. it should be criterion-referenced. Given this, it follows that different pupils may satisfy a given criterion at different ages: to tie the criteria to particular ages only would risk either limiting the very able, or giving the least able no reward, or both.

(TGAT, 1988, paragraph 99)

The report also emphasised that criterion-referencing in this manner would make it possible to monitor changes in national educational standards over time.

Ultimately, the goal of developing a tightly criterion-referenced assessment system proved to be very challenging to implement, and the model was radically loosened, if not entirely abandoned, during the mid-1990s (Dearing, 1993; Daugherty, 1995; Shorrocks-Taylor, 1999).

Scholarly roots

The roots of the CASLO approach are not straightforward to uncover. The architects of the new model – including Gilbert Jessup, of whom more will be said later – tended not to dwell on historical matters. Indeed, although the first major book on the adoption of the new model did include a chapter on its ‘background and origins’ (Tuxworth, 1989), it provided an oddly truncated account, which made it sound as though NVQs were little more than an adaptation of a North American approach to teacher education from the 1960s.

The roots of the CASLO approach were certainly North American, although they stretched back further into the first few decades of the 20th century. These roots were fundamentally educational, but it is interesting to note how they were pioneered by scholars who had a particular desire to improve assessment. The dominant theme here – from the Objectives Movement through to the Criterion-Referenced Measurement Movement – was the need to ensure that assessment is as comprehensive and authentic as possible, given the negative consequences associated with partial and inauthentic assessment. These concerns were just as salient in England as they were in the USA, which is why these movements migrated. In particular, there was a strong appetite in England for tackling over-reliance on the written exam format, widespread concern over the lack of attention to higher-level skills, and a strong appetite for addressing the prevalence of failure.

As the Objectives Movement began to influence educational thinking in England, particularly during the 1970s, there was a certain amount of resistance from scholars of education. This was a period prior to the introduction of the national curriculum, during which teacher control of the curriculum was hotly debated. It is easy to see how the idea of prespecifying educational outcomes might have appeared to embody one side of this debate, while many academic educationists continued to argue for retaining teacher autonomy, which was more consistent with the other side of the debate. Conversely, as the Criterion-Referenced Measurement Movement began to influence educational thinking in England, particularly during the 1980s, the idea of clarity over what students needed to learn and be assessed on seems to have been somewhat less controversial. Clarity would provide a necessary foundation for effective formative assessment as well as for improving summative assessment.

Wider sociopolitical context of the 1980s

As we have already noted, academic debate surrounding the introduction of NVQs, and the CASLO approach more generally, tended to focus on its conceptual basis. Williams & Raggatt (1998) chose instead to focus on the economic, institutional, and political factors that helped to explain the origins of these new competence-based vocational qualifications. From documentary analysis and interviews with policy

makers, officials, and consultants, they identified 4 interrelated beliefs that seemed to capture the underlying rationale for reform. Their analysis extends ideas from our earlier discussion of the 1970s landscape.⁵⁸

First, education, generally, had failed to prepare learners for the needs of employers and employment. The economic recession of the 1970s had focused attention on the degree to which education was preparing young people adequately for the world of work, particularly low-achieving students. Critics claimed that many students were leaving school essentially unemployable owing to a lack of basic skills, and with anti-industrial attitudes cultivated by the education system itself. Policy makers decided that this must change, and that education – particularly further education – should be refocused to respond to the needs of employers and employment. This became known as the ‘new vocationalism’ of the 1970s and 1980s, to some extent an expression of distrust of the educational establishment. This new policy stance led to the development of curriculum initiatives such as the Technical and Vocational Education Initiative (TVEI) during the early 1980s, which required co-operation with local industries to provide occupationally relevant school- and college-based programmes for 14 to 18-year-olds (Ainley, 1990; Stanton & Bailey, 2004).

Second, the employment market (and therefore skills requirements) had changed over time, yet education and training provision had failed to keep up with these changes. England had transitioned into a post-industrial age: some sectors were advancing rapidly (including services), while others were in decline (including manufacturing). Entrants to the new job market required new skills, including expertise in new technologies, as well as higher levels of skill. Yet, in many of the advancing sectors – retail, hotel and catering, and caring, for instance – levels of skill were low, with an absence of appropriate apprenticeships and qualification suites.

Third, the apprenticeship system of the 1980s was not consistently delivering the goods. It was predominantly focused on craft industries, and apprenticeship numbers were continuing to decline. We noted earlier that the quality of education and training delivered to apprentices was highly variable during the 1970s. Williams & Raggatt explained that apprenticeships were unknown in many advancing sectors, and were typified by restrictive practices. In short, the apprenticeship system needed to be disrupted rather than reinforced. The move towards competence-based vocational qualifications was therefore associated with an attempt to improve quality, to improve coverage, and to free the system from artificial barriers. Most importantly, apprenticeship should no longer be associated with serving time, but with acquiring competence.

Fourth, rising levels of youth unemployment raised new challenges and opportunities related to getting young people off the streets and into employment. The Manpower

⁵⁸ Further insights into this sociopolitical context are provided in Raggatt & Williams (1999, chapter 2).

Services Commission (MSC) was a non-departmental public body, responsible to the Department of Employment, with a remit to co-ordinate employment and training services, and the work of the Industrial Training Boards. Active from 1974 to 1988, it published a number of influential reports from the late-1970s onwards, which outlined a new conception of training standards that was to become the foundation upon which NVQs were built.

As unemployment rose during the 1970s, the focus of the MSC shifted towards the short-term needs, and then to the long-term needs, of the unemployed. Building upon the Labour-initiated Youth Opportunities Programme (YOP), the Conservative-initiated Youth Training Scheme (YTS) was introduced in 1983 to provide a programme of integrated education and training for school leavers. It comprised a 1-year programme of on-the-job training, designed as a new model of apprenticeship that was intended to replace apprenticeship-by-time-serving. Although originally intended to cater for both employed and unemployed young people, very few employed trainees were enrolled on the scheme (Ainley, 1990). It soon became seen as a low paid, low status, last resort option for young people, with no clearly defined objectives, no recognised certification, and low completion rates (Ainley, 1990).

There was clearly a pressing need to enhance the status of the YTS system, and formal certification was deemed critical to achieving this goal. As policy officials began to design the principles of a 2-year programme of education and training, the need for a review of the qualification system became increasingly apparent (Williams & Raggatt, 1998). The white paper 'Education and Training for Young People' (DES & DE, 1985) announced that all trainees should have the opportunity to work towards a recognised qualification, and set in train a Review of Vocational Qualifications. As we will soon see, this review led to the National Council for Vocational Qualifications, and to National Vocational Qualifications. In 1990, the YTS was replaced by Youth Training, which specified that all trainees must follow a training programme that leads to a Level 2 NVQ (Raggatt & Unwin, 1991).

Summary of the pre-history

National Vocational Qualifications – which we have identified as the first CASLO qualifications of national prominence in England – embodied an approach to qualification design that departed radically from traditional TVET qualifications, including Ordinary National Certificates and Diplomas, Higher National Certificates and Diplomas, and Craft Certificates. As NVQs began to be rolled out, critics began to compare them unfavourably with the traditional qualifications that they were replacing. In a radical critique, Smithers adopted exactly this strategy, arguing that “well-known and respected” qualifications were being replaced by low quality NVQs (Smithers, 1993, page 10 and section 6). Indeed, he took this one step further by strongly implying that the awarding organisations responsible for existing

qualifications were fundamentally opposed to the new NVQ model, but were unable to express their true feelings owing to pressure to buy into the new system (Smithers, 1993, paragraph 4.13).

In fact, the story is more complicated than this. The qualification systems that existed prior to the introduction of the NVQ framework were far from perfect. There would, undoubtedly, have been examples of qualification suites that operated very effectively within particular sectors for certain purposes. However, the huge diversity of TVET qualification provision during the 1960s and 1970s would be enough, in itself, to raise questions concerning quality across the board. There was certainly concern that the content of many existing qualifications was driven more by what teachers and trainers felt comfortable delivering than by what apprentices really needed to learn (Raggatt & Williams, 1999).

Furthermore, there were widely recognised, long-standing, problems with existing qualifications – related to wastage, retardation, and failure – which helped to dispose commentators towards greater reliance upon centre-based assessment. As well as helping to ensure that TVET qualifications were tailored to local needs, greater reliance upon continuous centre-based assessment might also help to improve the comprehensiveness and authenticity of these qualifications, enabling them to target the higher-level competencies that written exams had often failed to reach.

Beyond concerns related to the contents and processes associated with existing qualifications, it is important to remember that formal qualifications were often not required by employers, even to complete an apprenticeship. Apprenticeship was primarily a matter of serving time rather than acquiring competence. In fact, the Haslegrave report estimated that the majority of technicians in the workforce had no relevant qualification at all (Haslegrave, 1969). Furthermore, even well into the 1980s, there were large areas of the economy where no relevant qualifications existed, or where uptake was very low (Raggatt & Williams, 1999).

Turning from off-the-job qualifications to on-the-job training, the situation had improved since the 1964 Industrial Training Act, but it was far from perfect. At least partly because there was no formal certification of on-the-job training, quality varied widely both within and across sectors. The apprenticeship system of the 1960s and 1970s clearly had serious flaws (Peters, 1967; Oates, 2004; Fuller & Unwin, 2009; Mirza-Davies, 2015).

Finally, the economic downturn meant falling apprenticeship opportunities and rising unemployment. Very many young people had no opportunity to receive either on-the-job or off-the-job training. The Youth Training Scheme was a key part of the policy solution to this problem, and NVQs had to work in synergy with this initiative. Given this trajectory, the NVQ model is sometimes said to have: “evolved out of the need to validate work-based training programmes for young people” (Lester, 2011, page 206). Yet, although YTS certification might well have ended up as the focal problem

for NVQ designers, it is important to appreciate that NVQs were also designed with a variety of peripheral problems in mind, including all of those summarised above.

In the wake of the Haslegrave report, 2 new organisations – the TEC and the BEC – had begun to address many of these peripheral problems as part of their attempts to rationalise existing qualification systems. In particular, they had attempted to improve the authenticity and comprehensiveness of qualifications that served industry and commerce. A key part of their solution to this challenge was the adoption of insights and methods from the North American Objectives Movement. Thus, TEC and BEC awards were based on outcome-based qualification models, which clearly prefigured the CASLO approach and NVQs more specifically. When these organisations merged to form the BTEC, this outcome-based tradition continued, although qualification models evolved following their merger, and continued to evolve over time. The fact that the other principal awarding organisations were also experimenting with outcome-based qualification models during the 1970s indicates that hostility within the sector was clearly not as deep nor as wide as the Smithers critique appeared to imply.⁵⁹ Indeed, it seems reasonable to conclude that outcome-based approaches had become mainstream by the mid-1980s, both as a foundation for TVET training programmes and as a foundation for TVET qualifications.

⁵⁹ It is still true to say that the BTEC was more vocal in its opposition to aspects of the NVQ model than were the other awarding organisations (see Sharp, 1999, for instance). The Council developed its qualifications more in keeping with the BEC tradition than the TEC tradition, emphasising the centrality of integrated learning. Conversely, learning, per se, received little attention within the NVQ model (Cantor, et al, 1995). Moreover, as we shall soon see, the nature of the outcomes that were specified for NVQs, and the manner in which they were specified, was quite different from the approach adopted for BTEC awards, which proved to be a bone of contention. Yet, neither of these issues should detract from the fact that BEC, TEC, and BTEC awards were always based on outcome-based approaches to qualification design.

Chapter 3. Genesis

National Vocational Qualification framework regulations required that NVQs should be specified in terms of outcomes, with formal criteria for ascribing their acquisition, and with a requirement that all specified outcomes must be acquired. That is, they specified the CASLO approach in full. This chapter on the genesis of the CASLO approach describes the emergence of NVQs, including their design, implementation, and evolution. It also describes other key qualifications that came to adopt the CASLO approach, including General National Vocational Qualifications (GNVQs) and BTECs.

NVQs

In the following subsections we will explain how the NVQ model emerged, before considering how the first NVQs were received and, consequently, how they and the context within which they were situated changed over time.

Background

In April 1986, a system of National Vocational Qualifications was proposed by a Working Group chaired by the industrialist Oscar De Ville. It was originally conceived as a framework into which existing qualifications could be accredited. Ultimately, though, it became associated with a new approach to designing qualifications. The Working Group also recommended a National Council for Vocational Qualifications (NCVQ) to administer the new system.

The model that the NCVQ was soon to specify as a template for developing NVQs had its roots in research and development undertaken by the Manpower Services Commission. In December of 1981, the MSC had published 'A New Training Initiative: An Agenda for Action' (MSC, 1981b) on the same day as the government had published its white paper 'A New Training Initiative: A Programme for Action' (Raggatt & Williams, 1999). These reports promoted the idea of 'standards of a new kind' – an idea that was soon to be unpacked by research and development teams working at the MSC and, subsequently, at the NCVQ.⁶⁰

⁶⁰ The idea of 'standards of a new kind' was first mooted in the 1977 MSC report 'A Programme for Action – Training for Skills' (Raggatt, 1991). A response to criticism of the extant system of apprenticeship-by-time-serving, the new approach would be based on clearly defined and testable standards of occupational competence, which would help to improve access for both young people and adults. Plans were laid out in more detail in 'A New Training Initiative: A Consultative Document' (MSC, 1981a).

De Ville report

The 1985 white paper 'Education and Training for Young People' had concluded that future economic competitiveness depended on a coherent system for the assessment and certification of vocational competence. It set out proposals for a review of vocational qualifications in England and Wales. Under the chairmanship of Oscar De Ville, a Working Group was established by the MSC in conjunction with the DES. Its remit was to recommend a structure of vocational qualifications that would:

1. be relevant to the needs of people with a wide range of abilities
2. be comprehensible to users
3. be easy of access
4. recognise competence and capability in the application of knowledge and skill
5. provide opportunities for progression, including progression to higher education and professional qualifications
6. allow for the certification of education, training and work experience in an integrated programme

In his preface to the Review of Vocational Qualifications (RVQ) report, De Ville explained that:

In our recommendations we have sought to build on what is already good in present arrangements. There are many examples where local co-operation provides freshness, vitality and relevance. But nationally there is a lack of pattern or coherence; no clear overall accountability for vocational qualifications or ensuring standards; no assurance of progression or transferability. In spite of a plethora of institutions there are gaps. **In short there is no single focus for attention.**

(De Ville, 1986, Chairman's Preface)

The report began by highlighting a number of strengths and weaknesses of existing arrangements. Strengths included credibility, diversity, and partnership. Weaknesses included, opaqueness, duplication, gaps, inaccessibility, lack of take-up, and insufficient recognition of informal learning. On the positive side, the report identified "generally dependable assessment procedures and testing arrangements" whereas, on the negative side, it identified "assessment methods which are biased towards the testing of knowledge rather than skill or competence" (De Ville, 1986, page 1).

The report also reiterated 4 weaknesses that had been identified in the 1985 white paper, which concerned inadequate opportunities for:

1. individual achievement certified by one part of the system to be recognised by other parties or parts of the system

2. testing of skills and competence as well as knowledge and understanding
3. recognition of learning achieved outside formal education and training situations
4. flexible patterns of attendance and learning

In response, the report offered numerous recommendations, which included establishing:

- a National Council for Vocational Qualifications
- a national framework for vocational qualifications – the National Vocational Qualification framework
- objectives for a vocational qualifications system

The report anticipated that the NCVQ would provide a much-needed focus for a sprawling tripartite sector, which then comprised:

- examining and validating bodies (including City & Guilds, BTEC, RSA, LCCI, Pitman, and the REBs)
- examining and accrediting professional bodies (around 250, 76 with Royal Charters)
- industry training organisations (ITOs) and statutory testing facilities (including around 120 largely non-statutory ITOs and around 85 joint industry councils)

The NCVQ would be given a remit to secure a comprehensible system of relevant, credible, accessible, and cost-effective qualifications. Its primary function would be to exercise a quality assurance role by accrediting qualifications that had been developed by approved organisations to the NVQ framework.

De Ville envisaged that the NCVQ would bring coherence to the system by incorporating existing qualifications. This would certainly require changes to those qualifications. However, there was a strong sense that existing qualifications should be “brought within” the framework, rather than the NCVQ designing a “completely new structure to replace existing qualifications” (De Ville, 1986, page 25).

Yet, the Working Group also made many quite specific recommendations, which were interpreted in a manner that would ultimately frustrate this ‘onboarding’ presumption. These included recommendations for awarding credit, for credit accumulation across qualification components, and even credit transfer. Critically, De Ville insisted that:

A certificate that indicates performance in a written examination which tests the ability to describe, to state facts or to develop a logical argument is valuable but it is not a statement of competence as we would wish to have it. Many existing vocational qualifications are of this type and most fail to give recognition to work-based learning. Likewise a certificate that indicates the ability to exercise a skill or

to perform a limited and sometimes artificial task is useful, but it is not a statement of competence within our meaning. In satisfying the criteria for the National Vocational Qualification, accredited awards should not continue these deficiencies of most current certification.

(De Ville, 1986, page 33)

Instead, the Working Group believed that vocational qualifications should be defined as follows:

A vocational qualification is a statement of competence clearly relevant to work and intended to facilitate entry into, or progression in, employment, further education and training, issued by a recognised body to an individual. This statement of competence should incorporate the assessment of:

- skills to specified standards;
- relevant knowledge and understanding;
- the ability to use skills and to apply knowledge and understanding to the performance of relevant tasks.

(De Ville, 1986, page 17)

The report added that neither existing assessments of knowledge related to occupational skills, nor existing assessments of performance of skills, necessarily indicated occupational competence, by which the Working Group meant: “the ability to perform satisfactorily in an occupation or range of occupational tasks” (De Ville, 1986, page 30). To underpin the NVQ framework, the NCVQ would therefore need to work with sectoral standards-setting bodies to secure the specification of new standards of occupational competence.

It is worth pausing to reflect on the intended scope of recommendations from the Working Group. The report certainly reads as though it were proposing a general overhaul of TVET qualifications – from craft and technician qualifications, through to higher education and professional qualifications, as well as the more specific problem of YTS certification. Yet, based on interviews with key officials, Hargraves identified a more pragmatic perspective, which recognised that the new system was likely to be restricted mainly to qualifications within the influence of the Department of Employment, that is, the “operative and craft level awards of City & Guilds, RSA and the various industrial training boards” (Hargraves, 2000, page 294). In particular, he noted De Ville’s reluctant acceptance that BTEC technician awards were unlikely to be part of the new framework. The problem of how BTECs were to relate to the new framework “simmered below the surface” until the end of the 1980s before it “boiled over” into debates regarding the General National Vocational Qualification

(Hargraves, 2000, page 295). We will consider the evolution of GNVQs and BTECs later.

NCVQ

The 1986 white paper 'Education and Training – Working Together' endorsed recommendations from the Working Group and, in October 1986, the NCVQ was established as a non-departmental public body responsible to the Department of Employment. Its remit was to implement recommendations from the Working Group (Jessup, 1991).

From 1987, with support from the MSC, Lead Bodies were appointed to design the standards of occupational competence that lay at the heart of the new NVQ framework. The framework itself took shape very quickly and the first NVQs became available in 1987. The NCVQ began with 2 directorates: quality assurance and accreditation. In 1987, a research and development directorate was added, headed by Gilbert Jessup.

Gilbert Jessup

Jessup is a key figure in this account of the genesis of the CASLO approach. As Director of Research, Development and Information at the NCVQ, he became the principal architect of the NVQ model. He gained considerable experience with outcome-based qualifications during his time as an occupational psychologist in the Royal Air Force (Tim Oates, personal communication). He was subsequently appointed Chief Psychologist in the Work Research Unit of the Department of Employment where, incidentally, he and his wife, Helen, authored a book on 'Selection and Assessment at Work' (Jessup & Jessup, 1975). His next role was in the Manpower Services Commission where he worked on the Youth Training Scheme until he was invited by Margaret Levy to join the European Social Fund project team, which elevated his profile in relation to assessment and qualifications (Tim Oates, personal communication).

Jessup developed a blueprint for the NVQ model while working in the Quality Branch of the MSC. In the appendix of his landmark book – 'Outcomes: NVQs and the Emerging Model of Education and Training' (Jessup, 1991) – he reproduced a technical note from March 1985, which unpacked implications of the MSC (1981b) report. He began by revisiting the report's influential statement on standards:

At the heart of the initiative lie standards of a new kind. Such standards are essential for the following reasons:

- i modernization of skills training including apprenticeship can only be achieved if we can replace time serving by standards of training achievement and

ensure all those who reach such standards, by whatever route and whatever age, are recognized and accepted as competent;

- ii if all young people are to have access to basic training, they and employers will want to have a recognized record of skills, knowledge and experience gained and;
- iii if there is to be wider access to opportunities for adults, there must be a recognized system which allows the individual to build upon what he has and secure recognition for what he has gained to date.

(Jessup, 1985, reproduced in Jessup, 1991, page 166)

He then identified various implications for assessment and accreditation, which he believed followed from the idea of 'standards of a new kind', including the need to:

- accredit workplace learning, consistent with the emphasis on work-based learning within recent training schemes such as the YTS
- separate standards from training courses, to ensure that accreditation is accessible to all learners, regardless of the 'route' to that learning
- reference standards against job requirements, that is, to criterion-reference rather than norm-reference
- move from sample-based to comprehensive assessment, to make sure that individuals have acquired all of the specified standards
- specify standards in terms of both relevant activities (and the conditions under which they need to be performed) and criteria by which success will be judged
- specify overall requirements for competence in terms of relatively small units of activity, suggesting a modular structure for awards ⁶¹
- incorporate assessment arrangements that can accommodate all learners, regardless of the 'route' of their learning, bearing in mind that specifying relatively small units of activity tends to lend itself to continuous assessment

These stated implications clearly prefigure the core characteristics of all CASLO qualifications, including the explication of both learning outcomes and assessment criteria, as well as the mastery requirement. They also highlight additional characteristics that were to become more specifically associated with NVQs.

Jessup championed a movement that wanted to turn education and training on its head. Further education institutions had been widely criticised for being too inflexible,

⁶¹ The MSC had been influenced by work on modularisation and credit undertaken by the Council for National Academic Awards and by the Further Education Unit, which had been influenced by developments in the USA (Tim Oates, personal communication).

too dependent on young people for their clientele, too focused on teaching rather than learning, and “too acquiescent in accepting as holy writ syllabuses handed down by others” (Boffy, 1990, page 185). Despite detailed scrutiny by awarding organisation committees, there were some very outmoded practices enshrined in the syllabuses of certain courses (Cantor, Roberts, & Pratley, 1995). In addition, the recession had increased demand for full-time provision. Colleges responded by increasing opportunities for college-based simulation of real work experiences, but these opportunities were often inadequate. It led to a situation in which many students ended up obtaining the same paper qualification as past craftspeople would have done, but without a fraction of their practical experience (Cantor, et al, 1995).

Jessup questioned the established role of colleges in upskilling the nation. He believed that education and training had to be refocused. In particular, he argued that attention needed to be diverted from traditional syllabuses, courses, and training programmes – which specified ‘inputs’ to learning – towards standards of a new kind that specified ‘outcomes’ from learning (Jessup, 1990).

In a Foreword to Jessup’s landmark book, John Burke, Senior Fellow at the University of Sussex, emphasised the extent to which Jessup’s “personal contribution has shaped so many developments in the emerging model” (Jessup, 1991, page ix). It is clear from Jessup’s own writing that he was passionate about this new model. In a Foreword to a book edited 2 years earlier by Burke – titled ‘Competency Based Education and Training’ – Jessup explained that he found it “exhilarating” to be grappling with “fundamental questions on the way we learn and behave” (Burke, 1989, page x). He believed that many learners had failed to achieve their potential under the old input-driven model and he saw the new outcome-driven model as a revolutionary antidote. This was particularly true in relation to assessment:

We shall, I hope, see the demise of the last minute swotting of information soon to be forgotten for examinations. We shall not need to play those games in the future – games which few enjoy and where the majority finish up losers. Assessment will be open (the word ‘transparent’ is coming into vogue when what it meant is simply ‘explicit’ – able to be seen, rather than seen through). What is assessed and the standards of performance required are open to both the assessor and the candidate alike. Learners will be able to make judgements about their own performance which will have implications for their own learning. Self-assessment will become an important component in learning. It will also often contribute to and initiate assessment by tutors and supervisors.

(Jessup, 1991, page 135)

Along with new standards must go new forms of assessment, very different from sitting examinations. The model only works if assessment can cover all the things

we want people to learn (and, more important, what the learner wants to learn). It also only works if assessment is more friendly and facilitates learning rather than acting as a deterrent or just an obstacle to be overcome.

(Jessup, 1990, page 18)

Assessment practices such as sampling, providing a choice of questions and adopting pass marks of around 50 per cent, are all imports from an educational model of assessment, which have little place in the assessment of competence.

(Jessup, 1990, page 20)

The fact that this top-down revolution was openly critical of the educational establishment – and the fact that it was championed largely by the government's Employment Department rather than its Department of Education and Science – must have felt infuriating to many TVET scholars and teachers.⁶² Many book chapters and journal articles were written during the 1990s on the merits and (far more frequently) demerits of the NVQ outcome-based model. These academic debates were characterised by vigour, passion, hostility, and anger (Bates, 1995; Ecclestone, 1997; Hargraves, 2000).

Specifying the new standards

Influenced by the North American Objectives Movement, National Vocational Qualification standards would be derived from an analysis of outcomes rather than inputs. They would focus explicitly upon the competence certificated by the qualification, rather than doing so implicitly via lists of syllabus content. Yet, as clearcut as that might seem, considerable ambiguity remained over how best to model the nature of occupational competence at the heart of each NVQ.

Mansfield (1989) identified 6 new models of occupational competence in use in England, which had been developed by various agencies from the early- to mid-1980s (the FEU, the NCVQ, the MSC, and so on). He noted that the existing training structure tended to be based on a narrow, task-based view of competence and standards, whereas he wished to advocate a far broader one. The NCVQ agreed. This lent support to a new approach to specifying occupational standards, based on functional analysis.

⁶² Williams & Raggatt described the Department of Education and Science as a “minor ‘bit player’ in the development of vocational education and training during the 1980s” (Williams & Raggatt, 1999, page 83).

Functional analysis

The new approach that NCVQ came to promote – based upon functional analysis – was heavily influenced by David Mathews and Bob Mansfield. They were inspired by their experience of working on the ESF Core Skills Project during the early-1980s (which had been jointly funded by the European Social Fund and the MSC). This was a YTS initiative, which aimed to unpack the idea of ‘core skills’ within a model of work-based learning, with a particular focus upon describing generic workplace learning opportunities. The project team relied upon existing approaches to describing work activities – based on job and task analysis – whereby any particular job or task could be deconstructed into the set of discrete activities that comprised it. These approaches worked well in contexts where the work was highly routinised and procedural, which was certainly the case for many YTS trainees.

Mathews and Mansfield were unsatisfied with these existing approaches, however. They recognised that the world of work was changing and the labour market was moving towards a situation in which few (if any) jobs would remain highly routinised and procedural. They fully agreed with the central message of the New Training Initiative that Britain needed to develop a flexible, adaptable workforce. And they believed that this required a new way of thinking about competence. In 1985, as an antidote to the narrow, atomistic conception of competence implied by job and task analysis, they proposed a new Job Competence Model. This explicated a far broader, holistic conception, based upon 3 interrelated components of competence:

1. task skills (of the sort elaborated by task analysis)
2. task management skills (co-ordinating activities, solving problems)
3. role and job environment skills (working within the parameters of physical, interpersonal, organisational, or cultural constraints and expectations)

This model informed a new approach to describing the nature of work, based on functional analysis. Whereas job and task analysis focused squarely on activities, functional analysis focused instead upon the intended results of those activities. In other words, it focused on outcomes, whether products or processes. Furthermore, whereas job and task analysis failed to represent holistic aspects of competence, functional analysis aimed to represent those aspects explicitly, which followed from its fundamental principle of specifying a work role, that is, an occupational function.

The Job Competence Model was widely acknowledged for its influence on the development of NVQs (Jessup, 1990; Debling & Hallmark, 1990; NCVQ, 1995). Functional analysis was formally adopted as the basis for developing NVQ standards (Jessup, 1991; Mansfield & Mitchell, 1996). In 1988 and 1989, the Training Agency, which had assumed responsibility for developing those standards, published 6 ‘Guidance Notes’ under the heading of ‘Development of Assessable Standards for

National Certification', which explained this new conception of competence, and how it would be represented via functional analysis:

1. 'A Code of Practice and a Development Model' (TA, 1988a)
2. 'Developing Standards by Reference to Functions' (TA, 1989a)
3. 'The Definition of Competences and Performance Criteria' (TA, 1988b)
4. 'The Characteristics of Units of Competence' (TA, 1988c)
5. 'Assessment of Competence' (TA, 1989b)
6. 'Verification or Monitoring of Assessment Practices' (TA, 1989c)

From the outset, functional analysis was promoted as the best available method for developing NVQs, particularly because "other approaches do not fully reflect the broad concept of competence" (TA, 1989a, page 5). However, the Training Agency fully acknowledged that the approach was still being tested. Indeed, some 2 years later, Jessup also acknowledged that the technique was "still being developed" (Jessup, 1991, page 36). So, it is not actually straightforward to provide a definitive account of the nature of functional analysis that was supposed to underpin NVQ development. The following account is based upon that provided in 1996 by Mansfield & Mitchell in their book 'Towards a Competent Workforce'.

Functional analysis operates by analysing the functions that are carried out by an occupational sector, as a whole, before drilling down into particular roles. This is represented in Figure 3, which adapts illustrations from Mansfield & Mitchell (1996, page 95 and page 281). In this example, the key purpose of the occupational sector in question – construction – is to "establish, maintain and modify the use of the natural and built environment..." This can be disaggregated into a number of key areas, such as "formulate and implement strategies and policies..." These key areas can then be further disaggregated into key roles and functional units.

The basic question at the heart of the disaggregation process goes like this: in order to achieve the outcome described by the key area, what needs to be done? At higher levels of analysis, this might describe the concerted activity of a team, while at lower levels it would describe what individuals were expected to do. The first outcome in the example from Figure 3 answers the question of what needs to be done like this: "monitor and review environmental changes and need" – thus providing a specification for the first functional unit (A1.1) of the first key role (A1). This analytical process continues by specifying further key areas (key roles and functional units) until the key purpose has been exhausted.

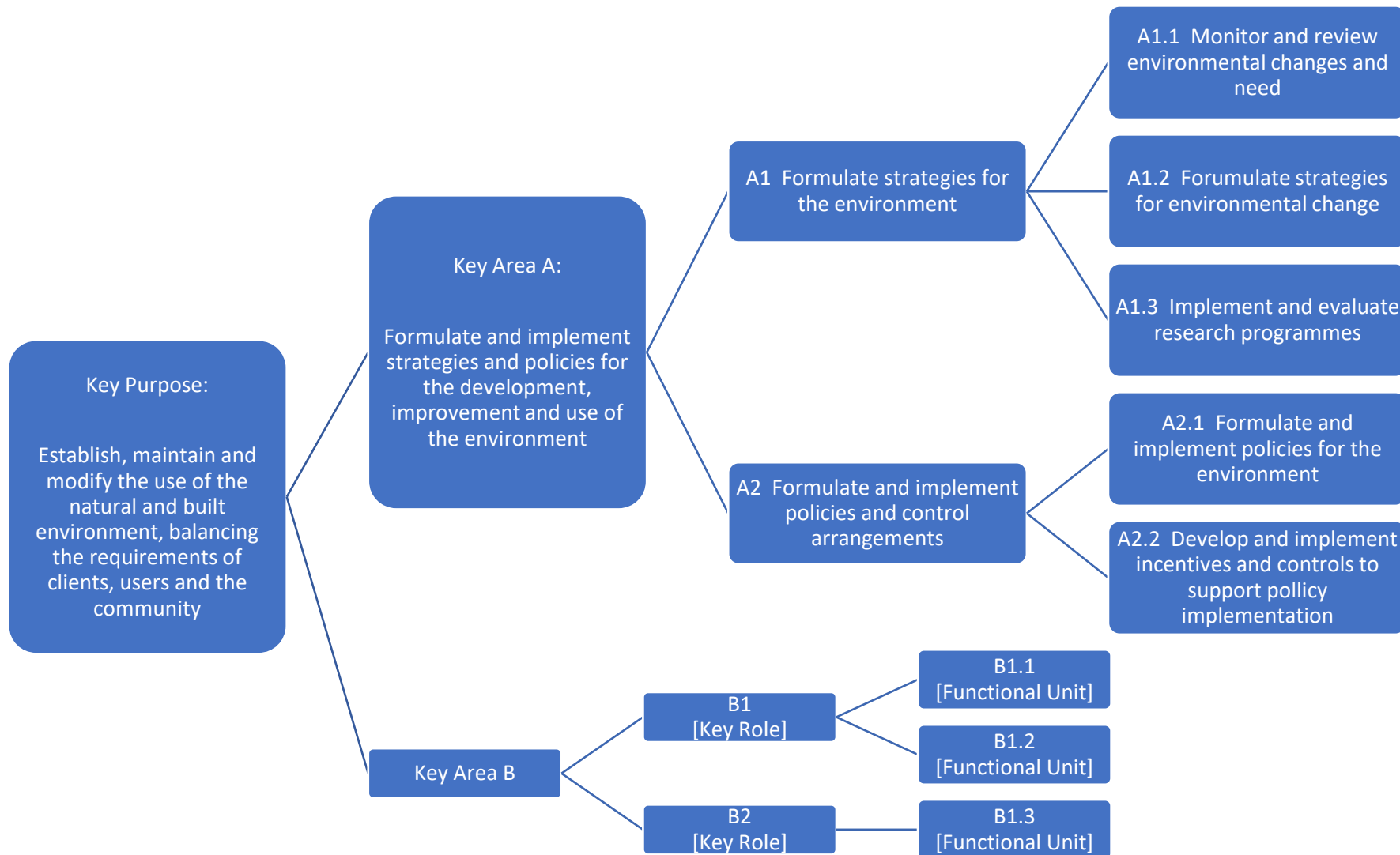


Figure 3. Illustration of a functional map derived from functional analysis

The idea of a key role takes us to the level of an individual worker, although this is an idealised role, not necessarily a blueprint for a specific job in the sector. This helps to clarify an important distinction between occupational standards and occupational qualifications as represented in Figure 4. Occupational qualifications relate to actual jobs, as opposed to idealised roles. Although some jobs may inherit their standards directly from functional units associated with a key role, others may not. Figure 4 (adapted from Mansfield & Mitchell, 1996, page 133) illustrates a situation in which an NVQ has been created for a job that draws its standards from multiple key roles.

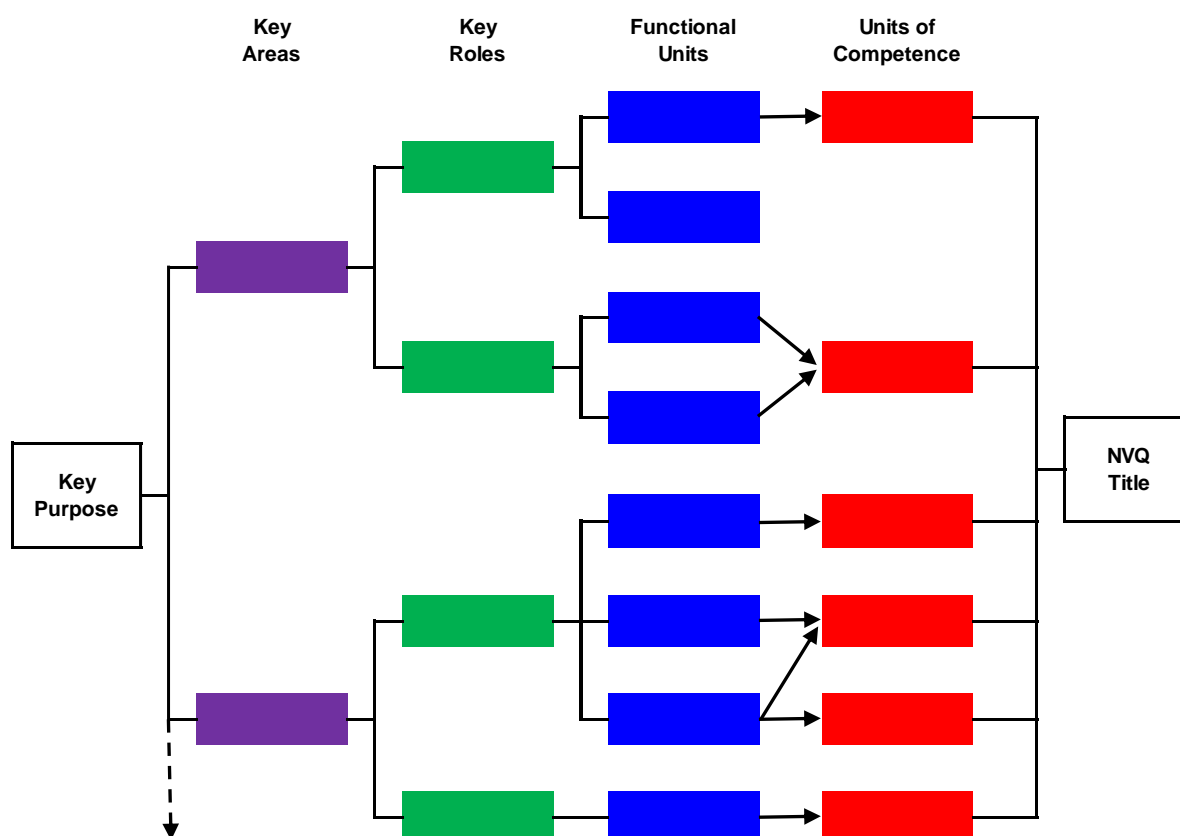


Figure 4. Relationship between occupational standards and qualification units

The principles of functional analysis also determine how functional units are specified (for occupational standards) and therefore how units of competence are specified (for qualification standards). From here on, we shall consider how units are specified using the qualification standards nomenclature, that is:

- qualification title – analogous to the key role
- unit of competence – analogous to a functional unit

As Mansfield and Mitchell describe functional analysis, it requires each unit of competence to be characterised in 2 dimensions, which involves specifying:

1. elements of competence (later to become known as 'learning outcomes'), and
2. performance criteria (later to become known as 'assessment criteria')⁶³

Each element of competence captures an outcome that needs to be achieved, specifying what needs to happen as the result of performing the role in question. For each element of competence, performance criteria capture the quality of performance expected, specifying how the role needs to be performed.⁶⁴

The following 2 examples are taken from Mansfield and Mitchell (1996, page 163) and derive from an occupational standard for the Plant, Animal and Land sector. They illustrate how units can be decomposed into elements of competence:

Unit 1: Monitor and coordinate the movement of people within sites

- E1.1 Welcome and receive visitors to the site
- E1.2 Care for visitors
- E1.3 Monitor and control unwelcome visitors

Unit 2: Commission, monitor and evaluate projects

- E2.1 Commission projects to enable objectives to be met
- E2.2 Monitor and evaluate the process and progress of projects against targets
- E2.3 Support project teams to enable them to achieve project objectives

Unit 1 indicates that an employee who is capable of monitoring and co-ordinating the movement of people within sites will be able to welcome and receive visitors to the site, care for visitors, and monitor and control unwelcome visitors. Note that these elements go beyond simply listing low-level, routine, procedural activities. E1.3, for instance, involves problem solving, that is, dealing with exceptions to the routine. E2.3 describes a high-level management function. Note also how E1.2 is fairly holistic, albeit perhaps fairly low-level, while E2.2 and E2.3 are both holistic and high-level. Mansfield & Mitchell (1996) suggest that there is an art to writing effective

⁶³ Later, we will see that a third dimension was subsequently specified for NVQs, to capture the expected range of application of competence.

⁶⁴ The nomenclature can be confusing, particularly when switching contexts from NOS to NVQ. In NOS nomenclature, elements of competence were nested within functional units. In this context, each element of competence was an occupational standard in its own right, as the terms were used interchangeably (Mansfield & Mitchell, 1996). In the qualification context, awards are often made at the unit level, which disposes us to think of the standard as a unit-level concept. Technically, though, NVQ standards were defined at the level of each individual element of competence, in terms of performance criteria.

standards: the scope of an element should not be so narrow that it begins to look like a specific task, but it should also not be so broad as to encompass distinct competencies that would be better elaborated separately.

It is particularly important that performance criteria are written carefully and with precision, to minimise ambiguity concerning the required performance standard. Typically, they describe a critical outcome or process plus an evaluative phrase. The following examples are taken from Mansfield & Mitchell (1996, page 163) and illustrate performance criteria written for qualitatively different kinds of outcome:

Physical products

Drawings and associated graphical material produced are complete, accurate and comply with design information and relevant documentation.

Interactive processes

Oral presentations are complete, accurate and presented in a pace, style and manner which are intended to maximise the trust and respect of all parties and are appropriate to the level of formality of the hearing.

Planned courses of action

The aims and objectives of production are identified in sufficient detail to allow planning to take place.

Process stages/requirements

Established conventions and procedures are followed.

Contingent outcomes which only occur if certain conditions apply

Incomplete and inconsistent input information is clarified promptly and appropriate and accurate amendments are made.

Mansfield & Mitchell were clear that performance criteria need to be as precise as possible without unnecessarily constraining action. However, they also noted that:

Most of the evaluative terms used in performance criteria require careful human judgement and consideration – and that is as it should be. The evaluative terms provide a benchmark which prompts participative discussion, negotiation and judgement – human attributes for a human system.

(Mansfield & Mitchell, 1996, page 192)

Governance

Fundamental to developing these ‘standards of a new kind’ was the principle that they must be employer-led. An organisation formerly known as the [Training Commission](#) – part of the Manpower Services Commission – was disbanded in September 1988. Its functions were absorbed into the Employment Department, and it was rebranded the Training Agency (TA). The Training Agency assumed responsibility for co-ordinating the development of National Occupational Standards (NOS) and NVQs. Each standard was developed by an employer-led representative body – an employer membership organisation, a professional body, or one of the many Industry Training Organisations that had been established since 1982 to keep training needs and standards under review (Debling, 1991; Laczik & Fettes, 2021).⁶⁵ These bodies were responsible for defining, piloting, and promulgating the standards.

Debling (1989) described how the 1988 white paper ‘Employment for the 1990s’ further delineated this process, by specifying that development would be spearheaded by Industry Lead Bodies. Since there would be no more than one set of standards per occupation or activity, there would also only be one Lead Body.⁶⁶ All new NVQs were to be approved by the relevant Lead Body, before final approval from the NCVQ (Gokulsing, Ainley, & Tysome, 1996).

Design

Substantial preparatory work at the MSC enabled the NCVQ and the TA to hit the ground running. This involved establishing the NVQ framework itself, specifying the NVQ design template, and promoting quality assurance processes.

Framework

A key objective for the NCVQ was to replace the array of widely divergent qualifications currently available to learners with a single, coherent, national system. This required the creation of a framework – the NVQ framework – which would specify the structure of this new system. The NCVQ represented this new structure using the diagram that is reproduced in Figure 5 (adapted from Jessup, 1990, page

⁶⁵ Consequently, the awarding organisations had no ownership of these standards and each National Occupational Standard was available to be incorporated within (equivalent) NVQs offered by multiple organisations. Wolf concluded that this led the ‘big 3’ vocational awarding organisations of the time – BTEC, City & Guilds, and RSA – to expand their activities, competing more overtly for trade instead of operating in a semi-monopolistic way (Wolf, 1995).

⁶⁶ By 1994, there were 160 Lead Bodies and 114 awarding organisations (FEFC, 1994a).

26). This suggested that NVQs would be available at multiple (although not necessarily all) levels across a range of sectors. Prospective NVQs would be accredited (by the NCVQ) within a particular sector at a particular level.⁶⁷

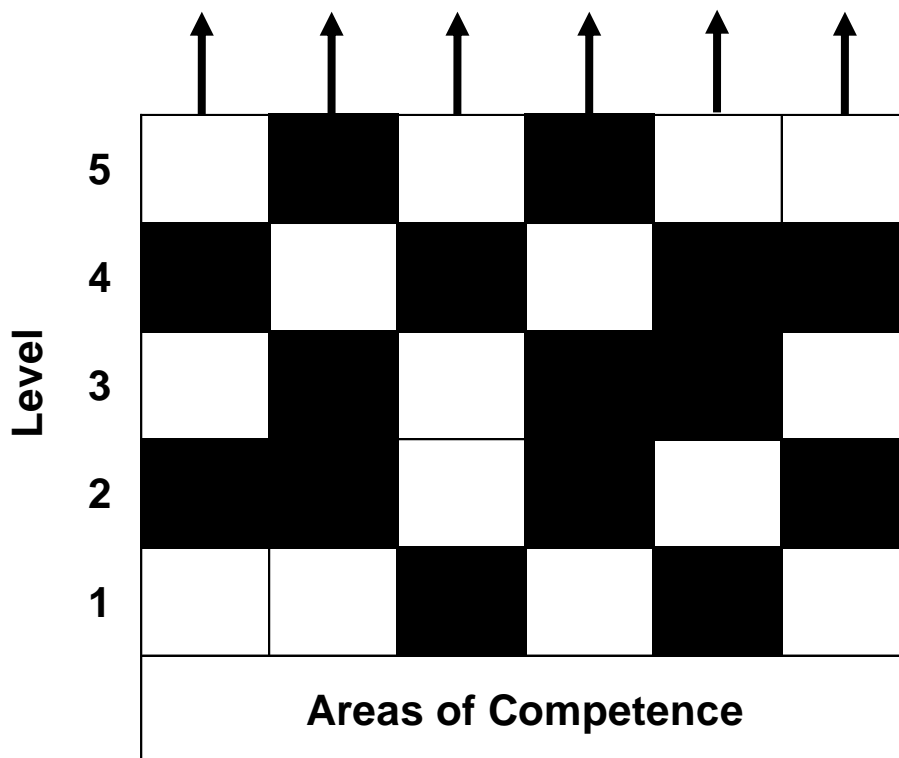


Figure 5. The NVQ Framework

Originally, only 4 levels were proposed, although this was soon raised to 5. Level 1 was associated with competence in activities that were mainly routine or predictable, or provided a broad foundation for progression. Level 4 was associated with competence in activities that were complex, technical, specialised, and professional, typically requiring a significant amount of autonomy and accountability. According to the white paper 'Education and Training for the 21st Century' (DES, DE, & WO, 1991) these levels would roughly map onto historical reference points as follows:

Level 5 – Professional Qualification, Middle Management

⁶⁷ Jessup (1991) identified 11 areas of competence, each subdivided into multiple subsectors. Tending animals, plants and land. Extracting natural resources. Constructing buildings, highways and related structures. Engineering. Manufacturing processes. Transporting. Distributing and selling. Providing leisure, accommodation and catering services. Providing health, social and protective services. Providing administrative and business support. Developing and managing human resources.

Level 4 – Higher Technician, Junior Management

Level 3 – Technician, Advanced Craft, Supervisor

Level 2 – Basic Craft Certificate

Level 1 – Semi-skilled

A core function of the NCVQ was to specify the criteria that qualifications were required to meet in order to be accredited to the new national framework. The first incarnation of 'The NVQ Criteria and Related Guidance' document was published in January 1988 (and revised in March 1989). It was originally envisaged that existing qualifications might be ported into the framework, albeit with some tweaking to make them fit. However, because the criteria ended up being highly specific, they effectively defined a new approach to qualification design. Thus, contrary to sentiments expressed in the De Ville report, the NVQ constituted a completely new type of qualification.

Design template

In contrast to traditional approaches, the NVQ design template was intentionally prescriptive in terms of (tightly) specifying outcomes and standards, yet also intentionally permissive in terms of (minimally) circumscribing implications for teaching, learning, and assessment. Indeed, flexibility came to be something of a watchword for delivering NVQs, particularly in relation to teaching, learning, and assessment.

The most significant feature of the NVQ model was the breadth of the construct that each NVQ was intended to represent, which was defined as nothing more nor less than the ability to perform an occupational role competently. This meant that each NVQ was designed to certify full occupational competence. It also meant that qualifications that were designed to certify knowledge, skill, or understanding beyond full occupational competence – for instance mathematical understanding that might be useful for progression but not for the current job – would not be accredited to the framework (Stanton & Bailey, 2004).⁶⁸

Competence

The foundation of each NVQ was a statement of competence. As noted earlier, this had 3 levels of detail:

1. the NVQ title

⁶⁸ This proved to be a particular challenge for the Business and Technician Education Council, which sought accreditation for many BTEC qualifications (Raggatt & Williams, 1999).

2. units of competence
3. elements of competence with associated performance criteria

A guiding principle of functional analysis was that statements of competence should have a common structure and grammar. This came to be viewed negatively, as jargon. Yet, NVQ designers deemed this to be essential for ensuring clarity (Mansfield & Mitchell, 1996). Thus, the NCVQ and the TA were very specific about how elements of competence should be written, insisting that each element should have an active verb, an object, and conditions. Jessup (1991, page 32) quoted an example from catering:

- maintain (active verb) standards of hygiene (object) in food preparation areas (conditions)
- assess (active verb) the physical condition of the patient (object) by inspection (conditions)

Performance criteria also had a common structure and grammar, which should always include both a critical outcome and an evaluative statement. This is an example from a Business Administration unit on filing from Marshall (1991):

- all materials are filed (critical outcome) without undue delay, in correct location and sequence (evaluative statement)
- all documents are classified (critical outcome) correctly (evaluative statement)

The NCVQ specified that elements of competence and performance criteria should:

- be stated with sufficient precision to allow unambiguous interpretation by different users, eg awarding bodies, assessors, trainers, and candidates;
- not be so detailed that they only relate to a specific task or job, employer or organisation, location or equipment.

(Jessup, 1991, page 17)

Range

Specifying statements of competence at just the right level of detail proved to be one of the most fundamental design challenges for NVQs. Research into criterion-referencing in the USA had already revealed the risks associated with wanting to make outcomes and criteria as unambiguous as possible. Wolf discussed this challenge at length, including the temptation for developers to be seduced into a “never-ending spiral of specification” resulting in standards that became unwieldy, unmanageable and, ultimately, unused (Wolf, 1995, page 55).

Element 1.1: Identify opportunities for improvement in services, products and systems

Performance criteria

- a) Relevant, valid, reliable information from various sources on developments in materials, equipment and technology is accessed and analysed for its significance at appropriate time intervals.
- b) Information on developments is disseminated to the appropriate people in a manner which is likely to promote its value.
- c) Information is related to current practices and used to identify opportunities for growth in operations and improvements in quality.
- d) Operations are continuously monitored and evaluated and where improvements can be made the necessary action is taken.
- e) Obstacles to change are accurately evaluated and measures to alleviate the problem implemented.
- f) Evaluation of the outcomes of previous developments is used for improvement.

Range indicators

Opportunities for improvement are identified:

- within the manager's line responsibility
- outside line responsibility, but where the manager has an impact.

Opportunities for improvement involve:

- personnel requirements/team composition
- employment/work practices
- work methods and patterns
- cost factors
- nature and availability of services and products
- quality of services and products
- methods to reduce waste
- new equipment/technology
- design of systems.

Implications of change are in terms of:

- profitability
- productivity
- quality of service/product
- environmental impact
- working conditions
- working relationships
- reactions of individual employees.

Analysis methods are

- qualitative
- quantitative.

Dissemination is to:

- higher level managers
- subordinates
- colleagues, specialists, staff in other departments.

Obstacles to change are:

- internal to the organisation
- external.

Figure 6a. Illustration of additional detail provided

Element 1.1: Identify opportunities for improvement in services, products and systems

Evidence required

Evidence must cover all those services, products and systems within the manager's line responsibility and those outside the line responsibility where the manager has an impact. Evidence must include the following items from the range:

- identification of the opportunities for improvement in:
 - personnel requirements/team composition
 - employment and work practices
 - work methods and patterns
 - costs
 - nature and availability of services and products
 - quality of service and products
 - methods to reduce waste
 - new equipment/technology
 - design of systems

[...]

Forms of evidence

Evidence can be outputs or products of performance, such as reports and documentation, supplemented by a personal report detailing actions that have been undertaken and why recommendations for action have been made. Evidence can also include extensive witness testimony from higher level managers, colleagues and subordinates.

In the absence of sufficient evidence from performance alone, questioning, projects and assignments based on real work situations may be used to elicit evidence of knowledge and understanding of the principles and methods relating to:

- accessing and analysing relevant information on changes to technology and resources
- analysing market need and marketing opportunities
- applying relevant items of legislation and organizational rules to actual/typical circumstances
- establishing, defining and reviewing objectives and performance measures
- informing and consulting others about problems and proposals
- monitoring resource utilization and costs and analysing efficiency and effectiveness

Figure 6b. Illustration of additional detail provided

Despite recognising this risk, the NCVQ and the TA decided, early on, that elements of competence would be open to different interpretations unless more detail was provided on what each element was supposed to cover. This led to the development of range statements, which indicated the variety of contexts across which each element was supposed to apply and, therefore, the variety of contexts across which each learner was expected to be competent. An NCVQ working group started exploring their use in 1988 (Wolf, 1990), and they became a formal requirement when the March 1991 revision of the NVQ criteria came into force (Gokulsing, et al, 1996).

Figure 6a presents an example of performance criteria for Element 1.1 of the 1992 Management Charter Initiative (MCI) Management Standard, supplemented by range statements, described here as ‘indicators’ of range (see Melton, 1997).⁶⁹ These were amplified further by a statement of ‘evidence required’ and a list of ‘forms of evidence’ (see Melton, 1997), both of which are illustrated in Figure 6b. This level of detail would have been specified for each element of competence within each unit.

Bearing in mind the requirement for generality in the specification of competence statements – such that they should not be defined in terms of particular task requirements, specific employer requirements, or suchlike – Jessup noted that range statements might indicate all sorts of dimensions of divergence, or variation, including organisation, equipment, materials, customers, products, and so on (Jessup, 1991). The idea, here, was that a competent individual should be able to perform the certificated functions in any company or job where they happened to practice, which emphasised the importance of being able to transfer competence across contexts. Range statements therefore played an important conceptual role in demarcating the boundaries of an element of competence – such that competence should not be assumed to transfer beyond the specified contexts – and provided a mechanism by which those boundaries could be redefined, if necessary, to reflect changing workplace demands (Mansfield & Mitchell, 1996). Having said that, Lead Bodies were instructed to be selective when constructing range statements and only include contexts that were common and critical.

Assessment

The logic of the NVQ model prescribed just one absolute assessment requirement: to assess all elements of competence (against their associated performance criteria) to determine when all of the competence requirements for the job in question had

⁶⁹ Mansfield & Mitchell (1996) drew a clear distinction between ‘range indicators’ (specified in occupational standards) that illustrate coverage for training purposes and ‘range statements’ (specified in qualification standards) that require coverage for assessment purposes.

been met. In other words, no sampling. Beyond this principled requirement, the way in which NVQ standards were defined – in terms of real-world performances – made other assessment decisions highly desirable or likely. For instance, it made sense for competence to be assessed in real-world environments, and for assessment to be a more continuous process, conducted hand-in-hand with learning. Unitisation of the statement of competence made unit-level certification seem quite natural too, with implications for assessment processes and timings.

Although the NVQ model was clearly oriented towards workplace assessment – and therefore to more naturalist assessment techniques including observation and discussion of naturally occurring events – it was open to a variety of assessment approaches. Not only was simulation a pragmatic fall-back option if naturally occurring events could not be observed, so too were many other assessment approaches. Mitchell (1989) argued that using multiple methods was likely to be important, as the evidence available from one method alone was unlikely to be sufficient for inferring competence. Having said that, she also indicated that certain methods were preferable to others, in the following order:

1. naturally occurring evidence
 - 1.1. ongoing work
 - 1.2. predetermined samples set in work place
2. specially elicited evidence
 - 2.1. performance (for example, traditional skills tests, college assessments)
 - 2.2. knowledge and understanding (for example, written or oral assessments)

More specifically, she suggested that an assessor ought to start by considering whether it was possible to assess the element of competence in question via a naturally occurring event. If not, then they should move down the list to the first viable alternative, even suggesting that an assessor could make do with a written assessment of competence if absolutely necessary. Again, though, the NVQ criteria were quite explicit in stating that performance should be demonstrated and assessed under conditions as close as possible to those under which it would normally be practiced. This also suggested that the most likely candidate for an assessor would be the learner's direct line manager or supervisor. If so, then this also meant that the assessor would be likely to be very familiar with the specified elements of competence and performance criteria being judged.

Knowledge and understanding

In January 1989, the TA hosted a symposium to consider the role of underpinning knowledge and understanding in competence-based vocational qualifications. Papers from the symposium were subsequently published in a booklet titled

'Knowledge and Competence' (Black & Wolf, 1990), with a Foreword from the Head of Standards Methodology at the Training Agency, Graham Debling, which read:

If good quality, cost effective vocational education and training is to be established we need a clearer insight into how and what knowledge and understanding underpins competence. [...]

This book seeks to contribute to a debate and investigation which is both almost as old as man and yet is only just beginning.

(Debling, 1990, page 2)

Confusion over the significance of knowledge and understanding within NVQs proved to be a huge destabilising influence for far too long. In a sense, protagonists like Debling were right that centuries of philosophical thinking had not yet furnished a straightforward account of knowledge and understanding, so perhaps practitioners could be forgiven for embarking on the NVQ project without a watertight account. Yet, heated, unrelenting debate over the nature of the NVQ competence model complicated rollout and undermined confidence in the new system.

The importance of knowledge and understanding to NVQs had never been in doubt. What remained unclear, however, was how, or even whether, NVQ standards ought to represent these constructs. According to both the original and revised editions of the NVQ criteria document (NCVQ, 1988; 1989), underpinning knowledge and understanding ought to be encompassed within the NVQ statement of competence itself. This could include specifying knowledge and understanding requirements as discrete elements of competence. Yet, within a year or so, following activities such as the TA symposium, the NCVQ had taken a far stronger line.

The 1991 NVQ criteria document (NCVQ, 1991) indicated that an NVQ statement of competence ought to be specified purely in terms of competence, leaving underpinning knowledge and understanding requirements implicit. This embodied the principle that knowledge and understanding were logically distinct from occupational competence (Jessup, 1991). They underpinned occupational competence – in the sense of being applied during the demonstration of competent performance – but they were not what was meant by competence. What was meant by occupational competence would be explicated by functional analysis, in terms of outcomes that a competent employee ought to be able to demonstrate to specified standards across a range of contexts.

Having clarified what competence means – in terms of what a competent employee is capable of doing – assessment ought to be a fairly straightforward matter of observing candidates, in real working environments, to determine whether they are capable or not. From this perspective, assessment ought:

- to target 'external' workplace achievements – which are achieved by a competent employee when performing their role – which means targeting specified activities
- not to target 'internal' cognitive achievements – which are acquired by a successful learner when learning how to perform their role – which means not targeting specified elements of knowledge or understanding

In theory, then, NVQ standards ought not to be defined in terms of knowledge and understanding, but in terms of competent performance. Likewise, in theory, observations of competent performance ought to be sufficient to infer competence, as long as they are sufficiently exhaustive (across contexts, over time, and so on). Indeed, according to Mansfield (1989), this was the only way to guarantee that an employee was genuinely competent.

Although the priority given to observing competent performance made sense in theory, it was unclear how strictly the principle could be respected in practice. A serious issue had already emerged following the move towards specifying range statements. Jessup acknowledged that practical constraints within real working environments often limited demonstrations of competence to single contexts, due to constraints on roles within particular organisations, or due to servicing only a limited range of clients, or suchlike. So, how might it be possible to determine whether a context-bound demonstration of competence was robust enough to transfer across contexts?

Well, the employee could be asked questions. For example, do they know how the features of successful performance will need to vary across contexts? Or, do they understand the underpinning theory of their role sufficiently well to be able to infer how their performance will need to vary across contexts? Competence can still be defined in terms of what a competent employee is capable of doing. But we can shortcut the assessment process by sampling competence in a small number of contexts and then use evidence of relevant knowledge and understanding to provide a warrant for generalising the inference of competence to the full range of contexts. It is important to acknowledge that this idea of supplementing observation with questioning was part of the original TA guidance:

We do not want to attribute competence until we can be confident that they will be able to perform to standard consistently, or across the required range of situations. So before attributing competence we normally need evidence of repeated demonstrations to standard, possibly in a range of different situations, and we may want to draw on more than one source of evidence by supplementing the demonstrations with questioning.

(TA, 1989b, page 5)

This approach was consistent with Mitchell's recommendations for triangulating multiple assessment methods (Mitchell, 1989). Jessup argued that it was necessary

to develop NVQs along these lines (Jessup, 1991). Mansfield & Mitchell (1991) developed a flow chart to help identify when performance evidence alone would be insufficient, as well as the kind of knowledge evidence that could bolster confidence in the attribution of competence. Wolf also agreed, but went a step further by suggesting that it was entirely legitimate to extend the definition of competence itself to include knowledge and understanding (Wolf, 1989).

The Employment Department published a second compendium in its 'Competence & Assessment' series, in 1992, which contained a chapter on assessing competence at higher levels. This recounted the experience of the Management Charter Initiative (MCI) in developing and promulgating standards for management, arguably the most complex area in which to specify competence according to the authors (Edmonds & Stuart, 1992). They began their analysis with a quotation from a recently published NCVQ guide to NVQs:

The opportunities for assessing performance will normally be limited in context or location ... and will not in themselves provide sufficient evidence relating to the full range of situations and contexts ... It is therefore necessary to supplement the assessment of performance with assessments of knowledge and understanding for most elements of competence. ... Assessment in NVQs requires evidence of competent performance, supplemented where necessary, by supporting evidence of underpinning knowledge and understanding. The importance of this is likely to be greater at higher levels.

(Reprinted in Edmonds & Stuart, 1992, page 49)

Owing to the particular challenges of assessing its management standards, guidance from the MCI proposed that there would not be a strong focus on direct observation. Instead, there would be an emphasis upon evidence arising from products, from witness testimony and, in particular, from personal reports. These reports might comprise written statements, or responses to oral questioning, which would focus on:

- details of actions taken
- reflections on actions taken, and
- knowledge of what was done and why

The MCI guidance emphasised the importance of employees being able to answer:

- 'why' questions (to explore understanding of underpinning principles), and
- 'what if' questions (to explore knowledge of variations across contexts)

Although, from very early on, it was accepted that it was often useful to assess underpinning knowledge and understanding independently of competent performance, the NVQ model (and the functional analysis methodology) still presumed that standards ought to be specified purely in terms of competence.

Having said that, the status of the 'NVQ model' itself was far from clear, as critical details evolved over time, including its stance on knowledge and understanding, as well as the addition of range statements. Indeed, it seems fair to say that theoretical details of the NVQ model were as much in flux, over the first few years of rollout, as practical ones. It is significant that Jessup's book included a chapter on 'The Problem of Knowledge' in a section entitled 'Outstanding Issues' (Jessup, 1991). It is also significant that Black & Wolf italicised the following extract from the first 'Guidance Note' penned by the Training Agency in 1988:

Each element of competence should describe something that a person who works in the particular occupational area should be able to do; an action, behaviour or outcome which the person should be able to demonstrate. *Or it should describe a knowledge or understanding which is essential in that it underpins sustained performance or facilitates the extension of skill to new situations within the occupation.*' (editor's italics).

(Black & Wolf, 1990, page 6)

Of the various contributors to the TA symposium, Bob Mansfield was the most 'fundamentalist' in his stance on not incorporating knowledge and understanding within NVQ standards (Black & Wolf, 1990). Yet, by the end of the year, it would appear that he had brought both the NCVQ and the TA on board.⁷⁰ Indeed, in an edited book in the same series as the Guidance Notes (Fennell, 1991), even Wolf respected this model as she described skills, knowledge and understanding as preconditions for competence, rather than as part of competence itself:

If standards are well and fully specified, they should assist the clarification of the knowledge and understanding implied by a unit or element of competence, both for learning and assessment purposes.

(Mitchell & Wolf, 1991, page 25)

The NVQ model

It is worth pausing to summarise details of the NVQ model, circa 1991 to 1992, following this early debate over the role of knowledge and understanding. First,

⁷⁰ It is worth noting that this debate over the role of underpinning knowledge and understanding was distinct from a broader debate concerning the breadth of the competence model, which had already been a site of conflict between the Manpower Services Commission and the Further Education Unit. The FEU was prepared to support MSC endeavours to move towards a competence-based system, but only on the understanding that education should have a central role, requiring a broader competence model: "The quid pro quo of this is a wider definition of competence than that associated with working life: embracing formal and informal learning, and extending beyond occupational skills into life skills." (FEU, 1984, i).

NVQs were based upon National Occupational Standards, which provided a comprehensive, outcome-based specification of an occupational role:

1. full occupational competence was broken down into component elements (the 'elements of competence' or 'learning outcomes' in more recent CASLO terminology)
2. further details were provided for each component element, to describe what performing competently looked like (the 'performance criteria' or 'assessment criteria' in more recent CASLO terminology)
3. the standards were intended to embody a broad definition of occupational competence, to describe what it meant to perform a role intelligently rather than mechanistically (hence the use of functional analysis rather than task analysis) ⁷¹
4. the standards were intended to be supplemented by a separate description of underpinning knowledge and understanding (to support the development and revision of NOS and NVQs, as well as to influence the development of effective learning content and processes) ⁷²

Second, NVQs were based on an authentic approach to assessment:

5. competence was intended to be inferred on the basis of evidence of consistently successful performance (the model distinguished clearly between performance that was observed and competence that was inferred, meaning that an isolated example of successful performance should be deemed insufficient, and evidence of successful performance should be required across a range of contexts)
6. assessors were expected to prioritise the most authentic assessment evidence available (ideally this would come from extended workplace assessment, although this was often supplemented by oral questioning, and was sometimes supplemented by other assessment formats including written testing)
7. evidence needed to be provided for all elements of competence (which is the 'mastery' requirement in CASLO terminology)

It is worth summarising these details because the original NVQ model has often been mischaracterised by critics. For instance, it has been said that:

⁷¹ Consequently, the elements of competence were intended to be general rather than highly specific, for example: "Edit existing text in a text processor." or "Write a report which evaluates potential solutions against known technical limitations and user's criteria." (both taken from TA, 1988b, page 6).

⁷² The place of knowledge and understanding in NVQ development is discussed thoroughly in Employment Department (1994).

- the NVQ competence model is inherently narrow and mechanistic⁷³
- the NVQ competence model denies the utility of constructs like knowledge and understanding, substituting any reference to knowledge and understanding with a reference to performance⁷⁴
- the NVQ delivery model eschews the very idea of a syllabus, course, or programme of learning⁷⁵

As we will soon see, the NVQ model was sometimes (perhaps often) implemented in ways that corresponded to these mischaracterisations. However, key details of the model were reasonably clearly articulated, even in early accounts by the Training Agency (1988a; 1988b; 1988c; 1989a; 1989b).

Flexibility

One of the selling points of the NVQ model was that it was designed to flex to the different circumstances that learners found themselves in.⁷⁶ Thus, NVQ certification was intended to be independent of:

- **mode of learning** (no expectation of having followed a particular course with a particular provider in a particular location – indeed, no expectation of having followed any formal course of learning at all – which was associated with the principle of accreditation of prior learning)

⁷³ For instance: “But, in general, we believe that ‘competence’ is the embodiment of a mechanistic, technically-oriented way of thinking which is normally inappropriate to the description of human action, or to the facilitation of the training of human beings.” (Ashworth & Saxton, 1990, page 24). Or: “More specifically, in the English NVQ system, competence is understood as the performance of a narrow set of tasks to a defined standard, and is thus bound to and reflects particular outputs.” (Brockmann, Clarke, & Winch, 2009, page 790). Or: “Since learning outcomes are, by their nature, narrowly conceived, what they measure is also narrowly conceived. It follows that there are difficulties in specifying learning outcomes for activities that, by their nature, are broad in scope, require underpinning knowledge for their performance and more complex personal characteristics than simple, visually observable, skills.” (Brockmann, Clarke & Winch, 2008, page 106).

⁷⁴ For instance: “All this confusion and equivocation seems to be the outcome of attempting to capture and describe, in behaviourist terms, something which is essentially non-behaviouristic, namely the development of knowledge and understanding.” (Hyland, 1993, page 61).

⁷⁵ For instance: “Gilbert Jessup, the chief architect of NVQs, proudly wrote of doing away with ‘the syllabuses, the courses or the training programmes [...]’ (Smithers, 1997, page 56).

⁷⁶ Note that the idea of ‘qualification flexibility’ described in this section is different from the idea of ‘workforce flexibility’ that had featured heavily in the New Training Initiative reports of the early-1980s, which argued that workers and companies of the future needed to be flexible and adaptable to cope with the uncertainties of a rapidly changing world. Mansfield & Mitchell (1996) argued that the key to workforce flexibility was effectively specified training standards, based on broad role specifications rather than narrow task specifications.

- **order of learning** (no expectation of having followed a particular learning trajectory, which facilitated flexible teaching across learners and cohorts)
- **age** (no minimum nor maximum age requirements, excluding legal ones)
- **minimum study time** (no requirement to have served as an apprentice for a specified amount of time)
- **maximum study time** (no requirement to have reached competence within a specified amount of time, which was facilitated by the ability to accumulate unit credits, as well as by the ability to transfer unit credits across providers)
- **session** (no expectation that either assessment or certification should be paced according to fixed calendar dates)

In other words, there should be no explicit nor implicit expectation that all learners ought to experience essentially the same regimented course of learning, prior to certification. Conversely, transparent certification requirements – laid out as elements of competence with performance criteria – should make it easier for learning experiences to be tailored to the particular needs of individual learners.

The potential for a more extended (ideally work-based) assessment process also enabled flexibility. Assessment tasks were not controlled centrally, meaning that the same outcome could potentially be assessed in different ways for different learners, depending on the circumstances of their learning and work. More generally, teaching and learning would not be constrained by practical requirements associated with external assessment, including timetabling.

Finally, the requirement to specify elements of competence and performance criteria as generally as possible – to apply with equal relevance across a broad range of occupational contexts – provided some flexibility for learning providers to adapt their provision to meet local or personal demands.⁷⁷ Thus, training for a national qualification could be reconciled with the desire to develop courses that were tailored to local and individual needs (Burke, 1989; Jessup, 1991).

Reconciling the idea of a national qualification with the divergent needs of local employers proved to be particularly challenging. Historically, England had prioritised diversity over coherence, to such an extent that City & Guilds was able to boast that there was no such thing as a typical qualification, because their qualifications were tailor-made to satisfy a defined need and category of industrial employee (Wheatley, 1959). Of course, 'bespoke provision' is simply the 'qualification jungle' by another name, which motivated the development of a single national system in the first place.

⁷⁷ It is worth noting that flexibility was also consistent with certain educational movements of the period, including the drive for increased personalisation of learning and campaigns for retaining teacher control of the curriculum.

So, there seems to be a sense in which the flexibility that was built into the NVQ model was an important concession to the very idea of a national qualification – a generic qualification not a bespoke one – that needed to possess a common currency despite divergent delivery contexts.

Ironically, despite flexibility frequently being cited as a selling point of NVQs, they also came to be criticised for their inflexibility. This tended to revolve around the detailed specification of learning outcomes in combination with the requirement that learners must achieve all specified learning outcomes for all units. This was problematic when standards were defined either too narrowly, where learners failed to acquire competencies that they actually needed, or too broadly, where learners were required to acquire competencies that they did not actually need (Debling, 1989; Field, 1995; Unwin, et al, 2004).

Quality assurance

A central argument in favour of the NVQ approach – when compared to a more classical approach to qualification design – was that it “demystified” assessment, because both outcomes and standards were stated clearly and comprehensively (Jessup, 1991, page 59). Transparency, so the argument went, builds validity into these qualifications. Wolf put it like this, albeit adding an important caveat:

As with all competence-based systems, the assumption has always been that assessment will be unproblematic because it simply involves comparing behaviour with the transparent ‘benchmark’ of the performance criteria. Unfortunately, in practice this turns out not to be the case.

(Wolf, 1995, page 64)

Although the NCVQ may well have oversold the transparency of NVQ standards and the validity of NVQ assessment, it never actually claimed that NVQ assessment would be unproblematic. Jessup, in particular, was quite open about the nature and scale of the challenges that would need to be faced in rolling out the new model (Jessup, 1989; 1990; 1991).

Particularly given how much flexibility the NVQ model afforded, the importance of establishing and following rigorous delivery processes could not be underestimated. Although a little late to the party, publication of the ‘The Awarding Bodies Common Accord’ (NCVQ, 1993) helped to address this challenge, significantly developing the original TA guidance on quality assuring NVQs (TA, 1989c). The main features of this Accord were:

- common terminology to describe the roles of individuals and organisations in the assessment and quality assurance system (including Assessor, Internal Verifier, External Verifier)

- certification to national standards for assessors and verifiers (units D32 to D35)
- defined roles in quality assurance for both awarding organisations and the centres they approve to offer NVQs (including Approved Centre, Awarding Body)
- explicit criteria for approving centres to offer NVQs (covering management systems, physical resources, staff resources, assessment, quality assurance and control, equal opportunities)
- quality assurance and control systems to ensure rigour and monitor equal opportunities implementation (including sample checking)

Implementation and evolution

It should already be clear that rollout of the NVQ model was far from straightforward. In fact, the very idea of a coherent rollout is misleading, as the NVQ model remained in flux throughout its early years. The following sections take up the NVQ story from the early 1990s, focusing on a number of key developments of particular relevance to the more general CASLO story.

Uptake

Having only been established in 1986, the NCVQ managed to accredit the first NVQs quite quickly, such that certificates were being awarded by 1988.⁷⁸ Figure 7 represents NCVQ data collated by Gokulsing, et al (1996, Appendix IV, page 87), which illustrate the number of certificates awarded from 1989 to 1993. It is clear that NVQs were largely catering for those working at Level 1 and Level 2. Across this period, the proportions of certificates awarded at different levels were 29% (Level 1), 58% (Level 2), 9% (Level 3), 5% (Level 4) and 0.2% (Level 5).⁷⁹

Figures from Field (1995) suggest that there were 1,346 NVQs in place by the end of 1993. Of these, 42% were at Level 2, although there appeared to be more NVQs at Level 3 (28%) than at Level 1 (16%), suggesting that certifications were higher, per qualification, at Level 1 than at Level 3. Field noted that these certifications were largely in the areas of goods and services, construction, and health care – rather than in engineering and manufacturing – and that the demand for Level 4 NVQs

⁷⁸ Given the speed with which the NVQ system was up-and-running, it is interesting to note that the NCVQ still ended up being criticised for how slowly NVQs were being made available during the early years (Sharp, 1999). Williams (1999) explained that early expectations concerning rapid rollout, including among policy makers, were based on the assumption that many existing qualifications could be ported straightforwardly into the new framework, which proved not to be the case.

⁷⁹ Different sources record different values for the numbers of certificates awarded, while still indicating the predominance of certifications at levels 1 and 2 (see, for example, Field, 1995).

related almost exclusively to accountancy (Association of Accounting Technicians awards in the main).

Field noted a powerful bias toward the mass purchase of low-skill entry level qualifications. The largest single market was among colleges offering full-time training to young people, where the Further Education Funding Council favoured NVQ uptake. The second major market was among providers of training for unemployed people, where funding also favoured NVQ uptake. Field characterised this as largely state-led rather than employer-led uptake.

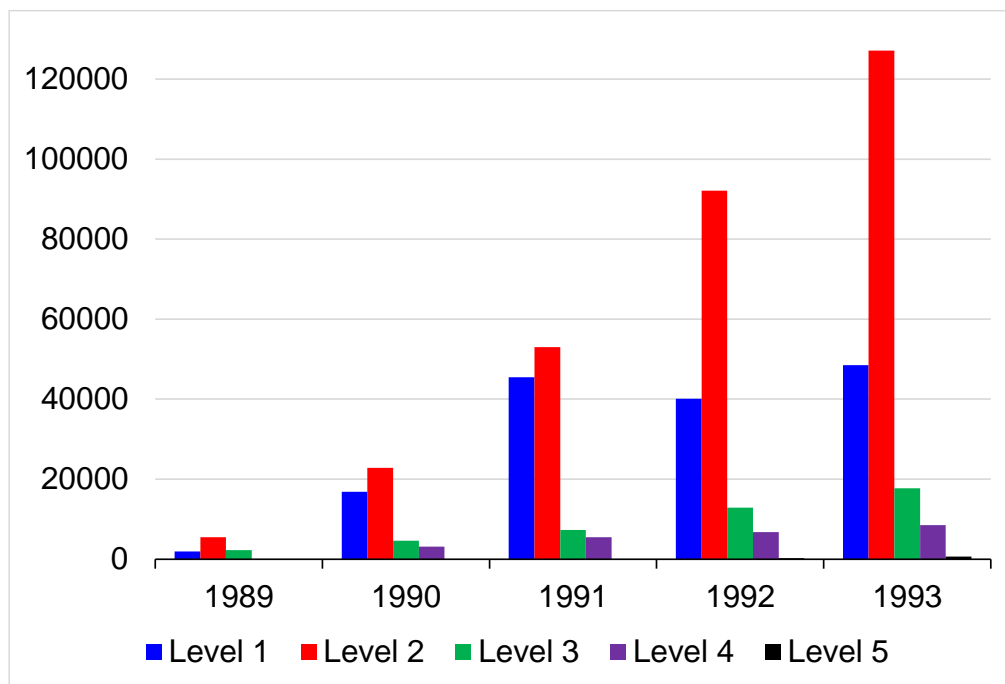


Figure 7. Number of NVQ certificates awarded 1989 to 1993

Shackleton & Walsh (1995) identified a variety of means that were used to suppress competing certificates, for example:

- requiring awarding organisations to phase out non-NVQ awards in related fields
- restricting central funding for training schemes and tax relief for individuals investing in their own training to the pursuit of NVQs
- requiring individuals who performed certain roles, such as health and safety, to have acquired relevant NVQs

Theory versus practice

To say that the NVQ model remained in flux throughout its early years oversimplifies the situation. It is not even that the model evolved through a succession of design

templates over time. Instead, at any particular point in time, NVQs with quite different design characteristics coexisted, for a variety of reasons.

This was partly attributable to an early compromise known as 'conditional accreditation' which meant that certain qualifications were awarded NVQ status despite only having been partially reformed, largely to ensure an income stream for the NCVQ (Williams, 1999). By October 1990, out of about 240 accredited NVQs, only a handful were fully accredited (Raggatt, 1991). Conditional accreditation helped the NCVQ to establish the system more rapidly, but it was a risky strategy and these pseudo NVQs accounted for much of the early criticism (Gokulsing, et al, 1996).⁸⁰

Beyond conditional accreditation, some organisations were granted permission to offer NVQs that departed substantially from the intended model in critical ways, including heavy reliance upon formal written exams. This continued over time, leading Young to conclude that bodies that occupied relatively powerful positions, such as the Association of Accounting Technicians, were able to shape the framework to suite their own needs rather than having to adapt to it (Young, 2011).

Divergence also occurred due to a lack of clarity and differences of opinion over the optimal approach for designing NVQs, which led to different organisations developing them in different ways. The Lead Bodies were provided with numerous 'Guidance Notes' on developing assessable standards for national certification, but they were also given considerable freedom to develop standards how they saw best, in consultation with whichever individuals or organisations they chose to collaborate (Debling, 1991). Furthermore, there was no statutory basis for requiring the Lead Bodies to comply with published guidelines (Ainley, 1990). Inevitably, some organisations developed standards far better than others, and some continued to plough existing furrows regardless of NCVQ or TA expectations (Ainley, 1990).

The recommended methodology also proved to be far from straightforward to apply. Standards that had supposedly been developed using (broad) functional analysis sometimes ended up looking like they had actually been developed using (narrow) task analysis. Arguably, the technical notes that Lead Bodies were required to follow when developing standards could be read in a way that appeared to justify the development of narrow, task-based competences (Raggatt & Williams, 1999).

Williams (1999) has argued that rollout moved increasingly towards a task-based orientation as NVQs catered increasingly for low-level jobs that reflected the impoverished content of many Youth Training Scheme programmes. Although NVQ

⁸⁰ The prevalence of external assessment within early NVQs is worthy of note in this respect. Steadman (1995) cited a study of the situation for NVQs as at October 1990. It indicated that almost all NVQs used supplementary evidence beyond observation of performance, and almost a quarter of unit assessments involved an externally set written test or exam.

standards were supposed to embody a broad definition of occupational competence – describing what it meant to perform a role intelligently rather than mechanistically – many NVQs failed to live up to that promise in practice.

Ultimately, the degree of mismatch between NVQ theory and NVQ practice makes it hard to judge the viability of the (intended) NVQ model on the basis of evidence from (actual) NVQ rollout.

Rollout

An early study conducted by Claire Callender, from the Institute of Manpower Studies at the University of Sussex, illustrated vividly many of the challenges encountered during the early years of implementation (Callender, 1992).

Commissioned by the Employment Department, the evaluation focused specifically upon the construction industry. This focus illustrated how successful implementation would inevitably have depended, at least to some extent, on employment and training structures within a particular sector. Rollout in the construction industry proved to be particularly challenging, given features such as the fragmentation of the sector (related to the increasing prevalence of subcontracting and self-employment), the large number of narrowly defined professional organisations with limited mutual understanding, volatility of demand and a highly mobile workforce, and a conservative approach to training with a strong focus on craft skills and time serving. Additional industry-specific challenges included increased capital and ongoing costs associated with the new NVQ model, such as having to adapt buildings to incorporate more authentic training activities, and higher costs associated with increased use of consumable training materials. On top of this, many employers, employees, and trainees remained unaware of NVQs.

More substantively, Callender identified issues that threatened implementation of the NVQ model itself. A critical concern was the lack of co-ordination and co-operation between Lead Bodies in the construction sector. This led to unnecessary duplication, but also to inappropriately narrow standards.

Callender was particularly concerned that the emphasis given to employer ownership had been at the expense of educational considerations. The Level 2 NVQs were too narrow, simplistic, and mechanistic. There was a lack of integration between NVQs at different levels, which ought to have been developed as a progressive sequence. On this specific point, she noted that the views of the construction industry and the Construction Industry Training Board – which were informed by practical and industrial relations considerations rather than pedagogical ones – were contrary to those of the Training Enterprise and Education Directorate of the Employment Department, the NCVQ, and City & Guilds (Callender, 1992).

Callender attributed problems of this sort partly to conflicting interests within the industry – with different interest groups attempting to ensure that their particular skill needs were met by the standards – but also to the fact that standards had been specified before the process of functional analysis for the construction industry had even been completed. Problems of this sort increasingly demotivated trainers, who had been very wary of change from the outset.

Other factors threatened the very idea of workplace assessment, which was at the heart of the new NVQ model. Construction trainees would not necessarily be exposed to all of the required elements of competence and range. It was often not possible to pause progress for the sake of assessment. NVQ standards were sometimes higher than those expected by certain employers. Costs were high. Paperwork was laborious. Supervisors were resistant to the idea of having to retrain as assessors. Supervisor bias and inconsistency were real threats. And so on. Ultimately, the very idea of workplace assessment was resisted, leading to a general consensus that assessment needed to be undertaken by training providers.

Attack

It is fair to say that there was a lot of criticism of both the NCVQ and of NVQs during the early years (Unwin, et al, 2004). One of the most high-profile critiques was mounted by Alan Smithers who had been commissioned by Channel 4 with the Gatsby Foundation to investigate the situation. His 1993 report – ‘All our futures: Britain’s education revolution’ – was promoted via a Channel 4 Dispatches television programme.

Focusing on both NVQs and the more recently introduced GNVQs, the report cited “real fears that there are deep flaws in the detail of what is being attempted” (page 8) and concerns over a “disaster of epic proportions” (page 3). Smithers traced the root of the problem to NCVQ insistence that students should be assessed “solely on what they can do rather than including also what they know and understand”. He characterised this as “behavioural psychology ruthlessly applied” and claimed that NVQs disdained knowledge, further insisting that “the notion of a syllabus is seen as antipathetic to the spirit of NCVQ” (all 3 quotations from page 9). The report contained a wide variety of criticisms of the NVQ approach and rollout, including the:

- lack of external testing
- flexibility promoted by the system
- lack of educational experience among Lead Body consultants
- narrowness of the standards
- incomprehensibility of the standards
- financial pressure on colleges to pass students

The report's recommendations included a number that were of particular significance to the CASLO approach:

That the content of NVQs and GNVQs should consist of an appropriate mix of skills, knowledge and understanding aimed at developing both vocational capability and educational achievement;

That in setting out the new content of NVQs and GNVQs, the schematic framework of "performance criteria" and "range statements" be superseded, and that course requirements be more simply and directly stated;

That the assessment of both NVQs and GNVQs should include written examinations as well as assessment of practical skills, independently set with marks externally verified;

(Smithers, 1993, page 43)

Gokulsing, et al (1996) described the Smithers report as an outlet for increasing demand for public debate over growing concerns with NVQs and GNVQs, with which the NCVQ appeared not to want to engage in public. Although its sensationalist reporting was prone to inaccuracy, bias and caricature – leading the NCVQ and even some of the people and organisations cited by Smithers to publicly denounce the report (Hodkinson & Issitt, 1995; Burke, 1995) – the attention that the report received led to further public scrutiny and reporting.

Inspection

The Further Education Funding Council was the inspectorate of its day. Its report on 'National Vocational Qualifications in the Further Education Sector in England' was, in effect, an extended response to the Smithers report, based on inspections during the 1993 to 1994 academic year.

The report claimed that strengths of the new learning programmes clearly outweighed any weakness for the majority of sessions observed (56%), a figure that was slightly better than for GCSE and slightly worse than for A level. It observed that the introduction of NVQs, with their emphasis on flexibility and responsiveness to individual students' needs, had led to a strong trend towards student-centred learning approaches.

Contrary to the claim from Smithers that providers were caving in to funding pressures, inspectors found "no evidence" of students being certificated as "having competences they did not possess". However, concern was expressed over "trainees' understanding of the principles underlying job competences" and "the poor levels of literacy and/or numeracy of some trainees" (all quotations from page 5). While acknowledging that more could be done to explicate underpinning knowledge

and understanding requirements, the report directly countered misleading claims in the Smithers report:

There is evidence of widespread misunderstanding of this view, and NVQs have been criticised for giving insufficient attention to knowledge acquisition. [...]

The proportion of time devoted to underpinning knowledge and understanding in the NVQ programmes inspected ranged between 15 per cent and 50 per cent, depending on the level and type of programme. Inspectors were generally satisfied with the level and quality of underpinning knowledge in terms of meeting NVQ requirements, although there were a few instances where it was deemed inadequate. There were concerns, however, about access to underpinning knowledge for some of the small proportion of candidates based in the workplace.

(FEFC, 1994a, page 16)

The report suggested that the NCVQ should insist upon greater clarification of the knowledge, understanding and core skills elements of NVQs prior to accreditation.

Beaumont report

In May 1994, the white paper 'Competitiveness: Helping business to win' announced a review of 100 of the most used NVQs and their Scottish counterparts (SVQs). Nearly a year later, former Chair of the Confederation of British Industry Training Committee, Gordon Beaumont, was invited to Chair the Evaluation Advisory Group that was to conduct this review, with support from the NCVQ and its Scottish counterpart (SCOTVEC). The review incorporated a wide range of research methods, including literature review, document analysis, stakeholder surveys, interviews, consultations and consultancy projects. Beaumont reported in January 1996 (Beaumont, 1996).

Although he recognised widespread criticism of NVQ implementation, Beaumont emphasised that the review had found widespread support for the NVQ concept, including the idea of competence-based standards, noting that he had seen such qualifications working effectively. He characterised those who questioned the concept as a "minority" (page 12) and linked their criticisms to early versions of the NVQ model, which had failed to pay due attention to knowledge and understanding.

That said, Beaumont did identify numerous significant criticisms of the NVQ model, which included (of most relevance to the CASLO approach):

- standards being "marred by complex, jargon ridden language" (page 13) that leaves candidates unsure of the competencies they are expected to acquire and that leaves assessors and verifiers unsure of the standards they are required to apply

- insufficient attention to core skills, which are key to enabling transferability
- prohibition (until recently) of optional units, which are key to ensuring relevance and that can prevent the proliferation of overlapping qualifications
- insufficient clarity over the centrality of knowledge and understanding requirements, typically where these are left implicit rather than stated explicitly
- the need to assess knowledge and understanding separately at higher levels
- excessive bureaucracy that leads to frustration, time wasting, unnecessary costs, and reduced uptake
- concern over lenient application of standards related to funding pressures, where funding requires completion within fixed time limits and assessors have a vested interest in timely completion

Although specifically asked to examine how external assessment might be included in NVQs, Beaumont identified mixed views and reached no clear conclusion other than that it is the “combination of methods which create rigour” (page 18) and that the NCVQ should therefore lay down the assessment methods appropriate to each situation with guidance on how to select them.

The most controversial recommendation from the Beaumont report was that ‘Part One’ NVQs should be developed. This was a response to ministers asking how knowledge and understanding might be separately certificated and whether this would be desirable. Beaumont noted that many employers already made use of traditional qualifications for this purpose. He suggested that greater use could be made of this approach if existing qualifications were made compatible with NVQs. Indeed, he formally recommended that traditional vocational and professional qualifications be made outcome-based and aligned to NVQs. Preparatory qualifications of this sort would also be useful to the many candidates who were not working. The controversial nature of this recommendation was captured by an article in the ‘Times Education Supplement Magazine’ following publication of the report:

John Hillier, chief executive of the National Council for Vocational Qualifications, said: “We will never change the concept of the NVQ through a preparatory qualification. The NVQ remains the goal.”

In sectors such as construction, he suggested, trainees might be sent to college by employers to pick up theory, then put their skills into practice in the workplace. The system would formalise existing moves by some employers to link with colleges for some elements of training.

The proposal for the so-called NVQ part one could reconcile the difficulty of applying one qualification for both employees and jobless school-leavers and adults.

Those without work might also take the preparatory qualifications, but then adequate Government funding would be needed to ensure they moved on to a full NVQ, said Mr Hillier. “To maroon a young person with only a preparatory qualification is not satisfactory.”

The decision on the part one qualifications would be left up to employers during consultation on the report, he said.

(Ward, 1996)

Revised Criteria

The revised NVQ ‘Criteria and Guidance’ (NCVQ, 1995) had actually been published a full year in advance of the Beaumont report. So, many of the problems identified in the report had already been addressed. Introducing the revised Criteria and Guidance, John Hillier stated that:

As a result, the document captures advances in thinking and methodology, without, however, changing the fundamentals on which NVQs are based. In particular, we have been able to reflect more fully the character of NVQs at higher levels and the role of knowledge and understanding within NVQs.

(NCVQ, 1995, page 2)

His foreword also recognised the new approach to designing NVQs around a mandatory core of units, with optional units tailored to particular employment needs. This flexible structure was to be matched by flexible assessment and delivery arrangements, which were adaptable to different organisational circumstances.

While attempting to remain true to the principles of functional analysis, and to the components of the Job Competence Model proposed by Mathews and Mansfield, the Criteria and Guidance document stated that it was “also necessary” to consider and “to reflect in the standards” the knowledge, understanding, practical and thinking skills, which are required for effective performance (quotations taken from page 17). Therefore, in addition to specifying outcomes (within elements of competence) that reflected the practical consequences of applying knowledge and understanding, the NVQ statement of competence now had to be accompanied by a formal knowledge specification. The document did not promote a particular approach to specifying underpinning knowledge and understanding, explaining that examples of good practice would be developed and disseminated subsequently.

The Criteria and Guidance document was also explicit over the breadth that NCVQ expected Lead Bodies to build into their specifications of standards. Consistent with a broad, role-based approach (rather than a narrow, task-based approach) it noted that:

People need to be able to communicate effectively with colleagues, organise and prioritise their work activities, respond to contingencies, make decisions, solve problems, apply ethical judgements, work safely and so on. It is the ability to integrate these demands when performing in the work environment that defines the competent individual. Lead bodies, therefore, are required to take a wide view of national standards which incorporates these broader aspects of competence.

(NCVQ, 1995, page 16)

Breadth was also promoted by the development of separately specified and assessed core skills units in communication, application of number, information technology, working with others, improving own learning and performance, and problem solving.

Finally, of particular relevance to the CASLO approach, the Criteria and Guidance document now specified evidence requirements in addition to the statement of competence. These requirements would indicate, on an element-by-element basis, the evidence required for a satisfactory judgement of competence, which might include types of evidence, methods of evidence gathering, and so on. The document also addressed the potential for confusing performance judgements with competence judgement by clarifying that:

It is not expected that any single item of evidence will be sufficient to establish competence in even the smallest assessable component of an NVQ, the element. Instead, it is expected that combinations of evidence should be used to attest to competence.

Combinations of evidence should be used flexibly to suit individual circumstances. Performance evidence should be combined with evidence of knowledge to cover the whole of the element specification, including range.

(NCVQ, 1995, page 29)

Rigour and responsiveness

The Qualifications and Curriculum Authority (QCA) came into being on 1 October 1997, assuming responsibilities previously assumed by the School Curriculum and Assessment Authority and the NCVQ. With a statutory remit to promote quality and coherence in education and training, it took the lead in designing and developing a new national qualifications framework (which we will consider later) and in ensuring clear and high standards across the system.

Recognising the need to enhance both the rigour of NVQs and their responsiveness to qualification users, QCA conducted a series of forums throughout 1998 and 1999 (QCA, 1999a). This provided an opportunity to discuss the new criteria, a new code of practice, and the development of risk management strategies, all intended to

enhance rigour.⁸¹ The QCA forums also provided an opportunity to consider how NVQs could be made more flexible and how awarding organisations could reduce bureaucracy.

QCA published its 'Arrangements for the Statutory Regulation of External Qualifications in England, Wales and Northern Ireland' in 2000 (QCA, 2000), which set out how both NVQs and NOS would be accredited to the new National Qualifications Framework, contingent on having met necessary accreditation criteria. In addition to criteria that were common across all qualifications, NVQs had to meet type-specific criteria, which were "designed to allow flexibility in the format of qualifications, while strengthening the processes to be followed, including greater emphasis on external quality control of assessment" (QCA, 2000, page 21). In particular, an assessment strategy had to be provided for each NVQ to explain:

how external quality control of assessment will be achieved, normally through the use of independent assessment. Where independent assessment is not recommended by the standards-setting body, other equally rigorous measures must be specified;

which aspects of the standards must always be assessed through performance in the workplace;

the extent to which simulated working conditions may be used to assess competence and any characteristics that simulation should have, including definitions of what would constitute a 'realistic working environment' for the qualification concerned;

the occupational expertise requirements for assessors and verifiers;

the amount and type of evidence to be collected.

(QCA, 2000, page 22)

Concern over the independence of assessors had been identified by the Beaumont report, particularly in relation to funding pressures, and guidance on independent assessment had already been prepared by the NCVQ (1997a). A revised version of this guidance explained that a significant part of the assessment ought to be carried out in a manner that was demonstrably independent of anyone with a vested interest in the decision (QCA, 1999b). This could be achieved in various ways, including via:

- externally set and marked tests or assignments

⁸¹ Note that Ron Dearing – in his 'Review of Qualifications for 16-19 Year Olds' (Dearing, 1996) – had echoed Gordon Beaumont's expectation that there "must be an over-riding requirement to demonstrate rigour" in NVQ assessment (Beaumont, 1996, page 19).

- visits from an external assessor
- externally set assignments, internally marked with external moderation

In addition to its Statutory Regulations, the QCA also published a 'Code of Practice' that was specific to NVQs (QCA, 2001). This built upon, and superseded, the Awarding Bodies Common Accord (NCVQ, 1993), with detailed requirements outlined in sections that included:

- Assessment and awarding
- Internal assessment
- Assuring quality in internal assessment
- External quality control of assessment including independent assessment
- Internal verification of internal assessment
- Assuring quality in internal verification
- Support and guidance
- Record keeping
- Awarding body quality assurance and control arrangements
- External verification of internal assessment
- External verification
- Sampling
- External verifier reports

These regulatory documents were supplemented by guidance including the 'External Verification of NVQs' (NCVQ, 1997b), the 'Internal Verification of NVQs' (QCA, 1998), and the 'Joint Awarding Body Guidance on Internal Verification of NVQs' (Joint Awarding Body Steering Group, 2001).

Raggatt & Williams noted that increasingly flexible assessment approaches – supported by the claim, from the Beaumont report, that rigour is underpinned by a combination of methods – were becoming a matter of concern to the NCVQ by 1997 (Raggatt & Williams, 1999). Flexibility of this sort – which was perceived as laxity by the NCVQ – risked throwing the baby out with the bathwater. The promotion of independent assessment, with its emphasis on tests and assignments, reflected a further departure from the original NVQ ethos. Raggatt & Williams noted that this “relaxation of the NVQ criteria continued under the auspices” of the QCA (Raggatt & Williams, 1999, page 163).

QCA eliminated the detailed prescriptions that NCVQ had placed on NOS. The new Statutory Regulations made no reference to 'functional analysis' and contained only minimal prescriptions, including the requirements that NOS should:

show the outcomes of competent performance, including the essential knowledge and understanding required;

be written in plain language and in a format which is easily understood by those who will use the standards;

(QCA, 2000, page 24 and page 25, respectively)

The 2004 revision of the Statutory Regulations specified requirements at a similarly high level, for example:

National Occupational Standards must: [...]

- describe the outcomes of competent performance; [...]
- include the essential knowledge and understanding required, the relevant technical, planning and problem-solving skills, the ability to work with others, the ability to apply knowledge and understanding, and other skills which will enhance flexibility in employment and opportunities for progression;

(QCA, 2004a, page 43)

Definitions in the final version of the 'NVQ Code of Practice' (QCA, 2006) also illustrated this more relaxed conception:

Competence: The ability to carry out activities to the standards required.

Content: The coverage of a qualification, programme, module, unit or other component, expressed as the knowledge, understanding, skills or area of competence that is covered.

(QCA, 2006, page 37)

An internal QCA report noted that this new flexibility to adapt the form and structure of NOS in response to sector-specific needs – which represented a rejection of strictures previously associated with functional analysis – had resulted in “a plethora of models” and new problems (QCA, undated, apparently circa 2001, page 11).

Apprenticeship reform

Ambiguity over the assessment of underpinning knowledge and understanding continued throughout the 2000s across numerous apprenticeship reforms. Apprenticeship reform had begun in 1994, as the publicly funded Modern Apprenticeship (MA) scheme was rolled out in response to continuing concerns over uptake, completion rates, and failure to secure employment (City & Guilds, 2014). Launched nationwide in 1995, the scheme was intended to revitalise the idea of apprenticeship training (Maguire, 1998) and to increase the number of young people achieving Level 3 NVQs, plus key skills, within 2 or 3 years. By the time of the

Dearing review, MAs covered two-thirds of the available NVQs at Level 3 (Dearing, 1996).⁸²

The importance of requiring more than just an NVQ (for example, key skills too) was reflected in the idea of an apprenticeship framework for achievement. Dearing made a number of recommendations concerning the new scheme, including this one related to the framework concept:

Employers should ensure that apprenticeships provide not only the necessary skills, but sufficient underpinning knowledge and understanding to enable Modern Apprentices, having obtained the NVQ level 3, to go on if they wish to part-time, full-time, or sandwich courses leading to diplomas and degrees.

(Dearing, 1996, page 40)

A few years later, this issue was developed in reports from the National Skills Task Force (NSTF), which had been appointed by Secretary of State for Education and Employment, David Blunkett, in 1997. Its final report presented a vision, goals, and main components for a National Skills Agenda. This included the proposal that apprenticeship “programmes at Levels 2, 3 and 4 should be available to all who want them and include key skills, assessed knowledge and understanding, and options for general education, so as to maximise transferability of skills and progression opportunities” (NSTF, 2000, page 7). This reinforced recommendations from interim reports for separate assessment of underpinning knowledge and understanding via related vocational qualifications (including existing BTECs). These qualifications would allow institutional providers to deliver vocational courses that were “more directly complementary to apprenticeship training” (NSTF, 2000, page 38). Government responded by proposing that Technical Certificates (TCs) would be developed to assess underpinning knowledge and understanding (DfEE, 2000).

Blunkett subsequently appointed a Modern Apprenticeship Advisory Committee to develop a 3-year action plan. Chaired by Sir John Cassels, the committee reported in September 2001 (Cassels, 2001). A lack of clarity concerning the “fundamental content” that every apprenticeship should contain led Cassels to propose a national framework for apprenticeship, which would specify required content and expected duration of apprenticeship at Foundation level (Level 2, age 16 to 19 entry) and Advanced level (Level 3, age 16 to 24 entry). The committee proposed minimum standards for key skills qualifications at both levels and offered reflections on the new TCs, including a recommendation not to reinvent the wheel where minor adaptations to well-established qualifications could fulfil the brief. It even proposed consulting higher education institutions to determine what content could be imported into TCs to facilitate progression from the Advanced apprenticeship. Cassels

⁸² Note that the ‘core skills’ nomenclature formally morphed into ‘key skills’ in 1997 (Mansfield, 2004).

recommended that a diploma should be awarded to recognise an apprentice's achievement of:

- an NVQ
- key skills qualifications
- a Technical Certificate, and
- any other awards gained during the apprenticeship

In early 2004, the QCA reported on an evaluation of TCs, conducted via a survey of colleges and training providers (QCA, 2004b). Since centres had first started teaching vocationally related qualifications designated as TCs in September 2002, approximately 200 qualifications had received this designation. The majority of respondents indicated that TCs were useful for assessing knowledge and understanding related to NVQs. Most TCs were also available for teaching to candidates who were not following the apprenticeship route, and around half of respondents adopted this approach. Some respondents were concerned about the use of exams in qualifications, particularly for students who had chosen the vocational route specifically to avoid them. QCA recommended that awarding organisations should review the use of external assessment in TCs, to ensure that its demand was relevant and suitable for target learners.

Later that year, the Sector Skills Development Agency (SSDA) published research that had been conducted with employer-led Sector Skills Councils (SSCs) to investigate whether apprenticeship delivery was meeting employers' needs. The report indicated an "overriding perception" that frameworks were "not sufficiently flexible or employer-centred" and straightjacket diversification (Pye, Pye, & Wisby, 2004, page 5). In effect, this challenged the very idea of a national apprenticeship framework. The inclusion of NVQs proved least controversial, although some whole sectors eschewed them entirely where other qualifications were better regarded. Employers in England particularly disliked the mandatory external testing of key skills. From a survey of 37 providers, 97% wanted the external testing ended and all of this group felt that key skills could more effectively be assessed via a portfolio approach. Many apprentices were simply failing to show up for the tests. More than half the SSCs disliked TCs for a variety of reasons, including the perceived irrelevance of content, duplication of learning and assessment with NVQs, problems of releasing apprentices for off-the-job training, and the inflexibility of Guided Learning Hours across certain subsectors. Where apprentices saw the NVQ as a gold standard, they often dropped out of the scheme without bothering to complete additional framework requirements.

Recommendations from the report implied a massive deregulation of the approach:

- allowing NVQs to be replaced with other vocational qualifications (or units)

- flexibility on the level of NVQ or vocational qualification required
- choice of whether or not to include a TC, or to replace it with a different programme of learning
- no longer externally testing key skills (deemed the most important request)
- continuing to regard apprenticeship as a scheme to be completed rather than a qualification to be achieved (contra the Cassels recommendation for a diploma)

The Leitch review of skills recognised widespread employer concern that the apprenticeship system was “complex and bureaucratic and often does not meet their needs” and that this contributed to low achievement rates (Leitch, 2006, page 98). Yet, Leitch was very positive about the role of apprenticeships going forward and recommended “dramatically” increasing the number of apprentices in the UK to 500,000 by 2020 (Leitch, 2006, page 21), as well as strengthening the role of employers. Government subsequently commissioned a review of all aspects of the apprenticeship system in England, which confirmed that apprenticeships would continue to play a central role (DIUS & DCSF, 2008). Plans to strengthen the apprenticeship framework included:

- improving the ‘blueprint’ to incorporate expectations of mentoring, progression, entry requirements and time off-workstation to train
- issuing a national completion certificate (rather than a diploma) at the end of the programme
- robust quality assurance against the revised blueprint
- integrating apprenticeship component qualifications within the new Qualifications and Credit Framework (QCF) to improve transferability and transparency
- greater employer ownership of apprenticeships

It was assumed that the QCF would be key to improving employer ownership:

As set out in the Leitch implementation plan, in future all vocational qualifications will be based on updated national occupational standards and will fall to be approved by Sector Skills Councils before being entered onto the QCF. This will provide a readymade bank of qualifications and units that employers, through their Sector Skills Councils, believe are needed in the workplace. In future, any organisation wishing to offer an Apprenticeship simply needs to submit to the relevant Sector Skills Council a short description of its plans, the qualifications and units it wishes to utilise, and how they meet the requirements of the strengthened Apprenticeships blueprint. In the case of an employer, the qualifications or units may include its own, accredited on to the QCF, so as to tailor the Apprenticeship to its own way of training.

(DIUS & DCSF, 2008, pages 36 to 7)

The integration of apprenticeship programme qualifications within the QCF is not incidental to the present report, as it was to become a statutory requirement that all QCF units should adopt the CASLO approach.⁸³ We will discuss the impact of QCF regulations later in this report.

The basic structure of the apprenticeship framework remained intact until the end of the Labour administration.⁸⁴ This included the following separately certificated components (see DIUS & DCSF, 2008; Skills Commission, 2009):

- a competency-based element, indicating the ability to carry out a certain occupation (typically certificated by an NVQ)
- a knowledge-based element, indicating theoretical knowledge underpinning a job in a certain occupation and industry (typically certificated by a TC)
- functional skills in numeracy and literacy (and other personal skills in some frameworks) – which replaced key skills and core skills
- a module on employment rights and responsibilities (often integrated within the TC)

Later in this report, we will consider how these requirements changed following the Richard review of apprenticeships.

Quality

Responsibility for developing National Occupational Standards originally fell to the Industry Training Organisations. Subsequently, ITOs, Lead Bodies, and Occupational Standards Councils were merged to form a smaller network of National Training Organisations. NTOs operated from 1997 until 2002, when they were replaced by an even smaller network of Sector Skills Councils. The Sector Skills Development Agency was established simultaneously, to fund, support, and monitor the performance of the SSCs (DfES, 2001). In April 2008, in the wake of the Leitch Review, the SSDA was replaced by the UK Commission for Employment and Skills (UKCES), which took over responsibility for overseeing the production of NOS by SSCs (and other standards setting bodies).

In the autumn of 2008, the UKCES commenced a consultation into whether NOS were fit for purpose. It concluded that the system needed to be improved, noting that the majority of employers still did not use them (UKCES, 2011a). The UKCES

⁸³ The regulator subsequently published 'Operating rules for using the term 'NVQ' in a QCF qualification title' (Ofqual, 2008b) to explain how the transition from NQF to QCF regulation ought to be managed.

⁸⁴ The 'Apprenticeship, Schools, Children and Learning' Act of 2009 led to the Specification of Apprenticeship Standards for England (SACE) in 2010, without radically affecting this structure.

proposed a strategy for change that involved improving NOS before promoting them more widely. A new set of quality criteria was central to this strategy, to ensure that:

high quality NOS, informed by a representative sample of employers, written in clear language and complying with common definitions are available for all significant functions in the workplace.

(UKCES, 2011a, page 11)

The resulting 'Quality Criteria' document (UKCES, 2011b) began by stating that NOS may only be developed by recognised bodies, that all personnel working on NOS in these bodies would need to be competent in their functions, and that any recognised body would be required to meet all of the specified quality criteria.

Significantly, the criteria required that NOS must be derived from functional analysis, and that all NOS should contain certain mandatory components: unique reference number, NOS title, NOS overview, performance criteria, specification of knowledge and understanding, and technical data. Certain optional components were also permitted: scope or range statements, values, behaviours, skills, and so on.

These requirements were elaborated in considerable detail in a 'Guide to Developing National Occupational Standards' (Carroll & Boutall, 2011), which was published by the UKCES alongside the criteria and strategy documents. The guide advocated a broad conception of competence, adapted from the Mathews and Mansfield Job Competence Model, and drew a clear distinction between occupational competence, *per se*, and the knowledge and skills that underpinned it. The guide focused upon the first 4 stages of developing and reviewing NOS: initial research, functional analysis, identification of existing NOS, and development of new NOS.

Although the guide was intended to promote quality, it is worth considering an example that the authors used to illustrate the articulation of performance criteria:

2R2/04 Deal with the arrival of customers

Performance Criteria

You must be able to:

1. Assist the customer to feel welcome in the hotel
2. Identify the customer's requirements
3. Ensure customer details are correct on the booking system
4. Offer alternatives for any services that are not available
5. Make sure the registration document is completed as required
6. Give accurate information to the customer about their room and its location

7. Promote the services and facilities of your organisation

(Carroll & Boutall, 2011, page 61)

In this example, the intended outcome is to 'deal with the arrival of customers' so the title describes the element of competence in question. Two points are worth noting concerning the articulation of performance criteria. First, each begins with a verb, such as 'assist' or 'ensure' which gives the impression of listing stages in a task. Second, the criteria do not always contain clear evaluative statements, without which the occupational standard remains unspecified (for instance, criterion 6 provides some clarity, while criterion 7 provides less). Mansfield & Mitchell (1996) identified both of these issues as common mistakes that are made when drafting performance criteria. Indeed, they identified both as typical of approaches that tend towards task analysis rather than functional analysis.

Norman Gealy (personal communication) has argued that the adoption of this approach in the UKCES guide, which had previously been advocated by the FEU, illustrates a turning point in the articulation of criteria – which generalised beyond NVQs to QCF qualifications – whereby criteria came increasingly to be written as mini learning outcomes, which meant that they no longer represented occupational standards. This, he believes, had a backwash impact on teaching and learning, such that mastering the domain became associated with having covered the necessary ground (a teaching expectation) rather than with having achieved the necessary standard (a learning expectation), which was precisely the problem that NVQs were originally designed to solve.

This is perhaps too harsh a criticism of the guide, *per se*.⁸⁵ The authors explicitly illustrated how their performance criteria looked quite different from a straightforward list of tasks. And they certainly did not ignore the evaluative dimension, as they wrote a full section on using evaluative words in performance criteria (although they did focus more on the risks of introducing ambiguity than the need to achieve clarity). The authors were grappling with a widely recognised challenge (post-Beaumont) that if employers and other users could not understand the NOS, then they were unlikely to make much use of them, and this affected the way that they wrote their performance criteria (Carroll & Boutall, 2011).

Finally, it is worth noting how knowledge and understanding were dealt with under the new quality criteria. Carroll & Boutall argued that knowledge and understanding requirements needed to be located within the NOS to indicate the breadth of

⁸⁵ It was certainly intended that: "performance criteria should be capable of distinguishing between satisfactory and unsatisfactory performance in the function covered by the NOS." (UKCES, 2011b, page 13).

competence expected by employers.⁸⁶ They provided the following illustration of requirements associated with the same element of competence described above (only the first 3 items have been reproduced):

2R2/04 Deal with the arrival of customers

Knowledge and Understanding

You must know and understand:

1. Expectations which customers may have when visiting the hotel, including standards of service
2. Why customers should be made to feel welcome in the hotel and the effect this has on their attitude to the business and the likelihood of repeat business and recommendations
3. Different techniques to help the customer feel welcome in the hotel and how to do this in different situations, for example when there are delays at reception or when there is a failure of equipment or services

(Carroll & Boutall, 2011, page 69)

These requirements were to be derived directly from the element of competence in question, and were only to include what was essential. They would typically list critical facts (the what), principles (the why), or methods (the how). Carroll & Boutall explained that these knowledge and understanding requirements had to be formally assessed, and that they tended to be assessed via questions, reflective accounts, or professional discussions.

According to this analysis, it appears that knowledge and understanding were not included in NOS as distinct outcomes in their own right with their own criteria. Instead, for each element of competence, they sat alongside the performance criteria, playing a different kind of supporting role. Consequently, these statements provided little insight into the level of knowledge and understanding required.⁸⁷ Presumably, where these knowledge and understanding requirements were incorporated within qualifications, such as Technical Certificates, the awarding

⁸⁶ Indeed, NOS were now defined as: "A statement of the standard of performance an individual must achieve when carrying out a function in the workplace, together with a specification of the underpinning knowledge and understanding." (UKCES, 2011b, page 27).

⁸⁷ The authors noted that: "If qualifications are developed from NOS, it is necessary to insert certain verbs such as 'explain', 'describe', 'list' etc., but this is not required for the NOS itself." (Carroll & Boutall, 2011, page 68). Yet, the simple addition of command verbs still does little to identify a boundary between having acquired the necessary level of knowledge and understanding and not having done so. This helps to bolster Gealy's argument.

organisations would have had to transform them into discrete learning outcomes with associated assessment criteria.

Stocktake

The NVQ story is central to our account of the origins and evolution of the CASLO approach in England, which is why we have recounted it in considerable detail. It is interesting and important for a number of reasons:

- the NVQ was the first CASLO qualification of national prominence
- it survived a long time, outliving its CASLO cousin, the GNVQ, by decades, which suggests that it must have got something right (or, at least, not entirely wrong)
- it continued to embrace the CASLO approach until its eventual (official) demise
- yet, it also remained controversial until its eventual (official) demise⁸⁸

In presenting this story, we have attempted to illustrate:

- how the CASLO approach was fundamental to the NVQ model
- how the NVQ model reflected a particular take on the CASLO approach, specifying additional features beyond the 3 core characteristics, some of which were critical to its fate (including its unique take on outcome specification)
- ongoing controversy that targeted both the model and its implementation
- continuing evolution of the NVQ model, while retaining the CASLO approach

This level of detail helps us to consider (if not definitively judge) the extent to which problems that beset NVQs might have been due to adopting the CASLO approach, per se, or to the particular version(s) of the approach adopted, to ancillary features of the NVQ model, or simply to poor implementation.

Principled design

Bearing in mind the amount of criticism levelled at NVQs over the past few decades, it is worth remembering that those responsible for designing the NVQ framework

⁸⁸ The regulatory framework that underpinned NVQs (including the 'Criteria for National Vocational Qualifications' – Ofqual, 2011a) was withdrawn in 2015 and remaining NVQs were then regulated under Ofqual's General Conditions of Regulation. Even prior to that, though, the NVQ system was gradually being dismantled, and many NVQs had already been replaced by or transformed into QCF qualifications. NVQ certifications in England fell from 77,580 in the 12 months up to quarter-4 of 2012 to 2,710 in the 12 months up to quarter-4 of 2015 (data taken from the [Ofqual Analytics](#) website). Ofqual still regulates a relatively small number of qualifications with 'NVQ' in the title, although NVQ is no longer recognised as a distinct qualification type.

believed that they were providing a principled solution to some very serious problems with Technical and Vocational Education and Training provision in England. Problems with off-the-job education and training included a predominance of overly theoretical qualifications, which were plagued by drop out and failure. Problems with on-the-job training included there being no guarantee that employees were achieving the right competencies to a satisfactory standard. These problems were compounded by high levels of unemployment, which raised new education and training challenges. The principled solution to these problems drew upon insights from North American educational movements, specifically the Objectives Movement and the Mastery Movement.

The objectives that the NCVQ located at the heart of the NVQ model were competence standards, which represented outcomes that would be manifested through competent performance in an occupational role. National Occupational Standards were therefore intended to specify occupational competence comprehensively and authentically, which would provide a solid foundation for subsequent curriculum, pedagogy, and assessment planning. They aimed to be authentic in terms of specifying the elements of competent performance that comprised occupational competence. They aimed to be comprehensive in terms of specifying all critical elements. By doing so explicitly, as a foundation for all subsequent planning, there would be no excuse for misaligned teaching or assessment.

The central innovation, then, was to refocus attention on what learners needed to learn, rather than on what teachers wanted to teach or what exam boards chose to examine. Just as importantly, learners would have to master all of the specified outcomes. As we will explain in more detail below, this was not just a certification requirement, but an educational expectation, consistent with insights from the Mastery Movement. This would be key to minimising drop out and failure.

Revolutionary zeal

Gilbert Jessup promoted his new model of education and training with revolutionary zeal. Critics certainly responded as though England was in the throes of a revolution, and commentators described exchanges between the NCVQ and its critics using war-like metaphors. Cantor described a situation in which the: “stage was thus set for some battles royal, with the control of further education curricula, and even the survival of colleges, as issues to be resolved” (Cantor, et al, 1995, page 60).

The impression of NVQ policy making as fundamentally confrontational was no doubt reinforced by the fact that it was being driven by the government’s Employment Department, rather than its Department of Education and Science. It seems likely that a consequence of this confrontational stance was to polarise stakeholders, which was particularly evident within the academic literature

(Hodkinson, 1992). This appears to have led to a situation in which certain mischaracterisations of the NVQ model became widespread, including the idea that it intentionally slayed certain sacred cows of the education world.

For instance, critics like Alan Smithers promoted the idea that the NVQ model was fundamentally opposed to the development of syllabuses and learning programmes (Smithers, 1993; Smithers, 1997). This mischaracterisation was fuelled by the manner in which outcomes were specified in the NVQ model, as competence standards, which gave the impression that the model had no place for knowledge and understanding. This, in turn, seemed to question the role of colleges in delivering NVQs, with their traditional focus on theory rather than practice.

It is true that the NVQ model was intended as an antidote to problems identified with extant qualifications. These qualifications had traditionally catered for trainees and apprentices who needed off-the-job education and training, with a focus on theoretical foundations, and they had often failed to fulfil even this limited role effectively. NVQs were intended to certify full occupational competence – not just the knowledge and understanding that comes from studying a book – and the principal yardstick of occupational competence was to be competent performance of an occupational role. This meant that workplace assessment would occupy a central position in the NVQ model, which again seemed to question the role of colleges in delivering NVQs.

Despite this impression, it is not actually true that the NVQ model negated the role of colleges. Instead, it implied that their role would need to change (see, for example, Nash, 1995; Stanton, 1995). More fundamentally, the NVQ model did not eschew the very idea of a syllabus or learning programme. Those who designed the NVQ model were very clear about this, for instance:

It is a primary focus of the new system that learning programmes, qualifications and assessment systems will be derived from clear and precise occupational standards, rather than standards being embedded unstated as a feature of qualifications, or within the processes of learning and assessment.

(Mansfield, 1991, page 12)

The point of adopting an outcome-based approach to qualification design was to focus attention on the proper foundation for planning curriculum, pedagogy, and assessment, that is, to focus attention on the outcomes that a qualification would need to certify if it were to serve its purposes adequately. In other words, outcomes were never intended to substitute for syllabuses or learning programmes. They were intended to provide a foundation for them. Exactly the same logic applied to determining the nature and scope of the knowledge and understanding that underpinned competence (that would need to be built into a learning programme) that was supposed to be inferred from detailed scrutiny of the intended outcomes.

Having said that, it is important to acknowledge a fundamental tension between NVQ theory and practice. Although the preceding analysis follows directly from the underpinning logic of the NVQ model – and although it was articulated explicitly by Mansfield (1989) and by others – it is still fair to say that NVQ rollout focused principally on assessment, and the importance of developing coherent learning programmes was often overlooked. Jessup insisted that the system needed to pivot to focus upon what gets learnt, not what gets taught, particularly as different learners might need different learning programmes. In practice, however, this refocusing shifted attention away from learning toward assessment, resulting in a system that was dominated by the voice of the assessor rather than by the voice of the teacher or trainer. In the introduction to this report, we explained our intention to explore whether problems that beset the CASLO approach were best understood as inevitable consequences of an unworkable model or as avoidable consequences from poor implementation. This would seem to be a good example of the latter. The very idea of an outcome-based approach is that it ought to provide a solid foundation for teaching, learning, and assessment. So, the fact that NVQ rollout was often associated with impoverished teaching and learning is unfortunate, if not ironic, but it was not inevitable.

Turning to the issue of underpinning knowledge and understanding – and bearing in mind that NVQ designers never denied the importance of these concepts to competent performance – it remains unclear why knowledge and understanding requirements were originally left implicit, rather than being spelt out explicitly, as they later came to be. At least in theory, the NCVQ could have required awarding organisations to be far more proactive in terms of NVQ syllabus development. Ainley observed that leaving so much work to educators and trainers alienated colleges and training providers (Ainley, 1990).

From a sociological perspective, Young (2008) has argued that NVQs were explicitly designed with a view to wresting control of the curriculum from providers (colleges) and placing it in the hands of users (employers). Consequently, NVQ designers were required to “avoid allowing the traditional syllabus-based approach to knowledge to return” to prevent colleges from reclaiming the curriculum (Young, 2008, page 141). This explanation might be overly simplistic.⁸⁹ But it does, once again, capture the idea of revolutionary policy making that unhelpfully polarised stakeholders, which seems undeniable.

⁸⁹ After all, colleges were not really in control of the curriculum before the introduction of NVQs, as syllabuses were traditionally specified by awarding organisations in partnership with multiple stakeholders, including employers as well as colleges. Indeed, even when empowered by the course validation model of TEC, BEC, and later BTEC awards, colleges often preferred to use centrally developed units (Cantor, et al, 1995).

It seems likely that part of the reason why the NVQ model appeared to eschew traditional college-based provision was the expectation that NVQs should be capable of servicing programmes like the Youth Training Scheme, which were designed to provide on-the-job training, often for fairly low-attaining young people within fairly low-level jobs (Hargraves, 1995). Many NVQs were developed for circumstances of this sort, where the need to unpack detailed underpinning knowledge and understanding requirements may have been less evident. And, of course, a fundamental tenet of the NVQ model was that it did not matter whether underpinning knowledge and understanding was acquired in college, in work, or anywhere else, just as long as it had been acquired.

The NVQ model can be seen as an extension – or perhaps even the culmination – of the approach pioneered by the Industry Training Boards during the 1960s. They, too, adopted an outcome-based approach to specifying training needs, albeit still recognising the traditional distinction between on-the-job (work-based) training and off-the-job (college-based) education. The NCVQ took this approach a step further by attempting to develop an even more comprehensive and authentic model of occupational competence, without having much at all to say about either on-the-job training or off-the-job education. However, its prioritisation of workplace assessment would certainly have appeared to have downplayed the importance of further education colleges. And the model itself, which lionised the concept of personalised learning programmes, would certainly have been challenging for colleges to implement. It seems fair to conclude that the NCVQ attempted to transform the concept of a TVET qualification from being focused squarely on education and colleges (Craft Certificates, ONCs, HNDs, and so on) to being focused squarely on training and employers (NVQs). In retrospect, this strikes us as an attempt to swing the pendulum of change from one inappropriate extreme to another.

Extreme critique

The theoretical basis of the NVQ model was heavily critiqued by scholars working in the TVET field, many of whom claimed that it was naïve. Although the present report is not the place for an extended evaluation of this literature, it seems fair to conclude that some of the most extreme criticisms were overstated, perhaps as a consequence of the kind of polarisation just discussed. This includes the claim that the NVQ model was behaviourist on at least 2 counts, and therefore fundamentally flawed.⁹⁰

⁹⁰ The suggestion that the NVQ model was 'behaviourist' seems to have become a matter of TVET dogma over the decades. See, for instance: Field (1991); Marshall (1991); Norris (1991); Ashworth (1992); Hodkinson (1992); Hyland (1993); Jones & Moore (1995); Tarrant (2000); Elliott (2001); Grugulis (2002); Halliday (2004); James (2006); Brockmann, et al (2009); Wheelahan, 2016; Murtonen, Gruber, & Lehtinen (2017).

First, proponents of the NVQ model were said to have been influenced by behaviourist philosophy, causing them to reduce knowledge and understanding to nothing more than competent performance. In fact, the concepts of knowledge and understanding were not reduced to competent performance within the NVQ model. It is true that occupational competence was explicated in terms of competent performance. But knowledge and understanding were defined quite differently – and separately – as constructs that underpinned occupational competence. The model even acknowledged that knowledge and understanding sometimes needed to be assessed ‘directly’ to provide ‘indirect’ evidence of occupational competence. This is not philosophical behaviourism.

Second, proponents of the NVQ model were said to have been influenced by behaviourist learning theory, causing them to specify occupational competence in a manner compatible with behaviourist learning approaches, such as stimulus-response conditioning. Consider the following claim, for instance:

This procedure clearly draws upon the work of the classical behavioural school of psychology. The work of Watson, Guthrie, Thorndike and Skinner is strongly represented. The classical behaviourists also concentrated upon the outcome of learning and judged the success of learning entirely by the behavioural outcome. This simplistic view of learning is now only of historical interest. It is surprising that the NCVQ have based so much of their work in this orthodoxy because the theoretical consideration of learning has advanced considerably in the last twenty or more years. Even the most radical behavioural psychologist would not now subscribe to the traditional view of learning so evident in the work of the NCVQ.

(Marshall, 1991, page 61, footnote references removed)

As noted earlier, it certainly is true that behaviourists, including Thorndike and Skinner recommended specifying intended learning outcomes in terms of behavioural objectives. Yet, non-behaviourists, like Tyler, also recommended this! Indeed, Tyler railed against the specificity of behaviourist objectives, citing for example a case from Thorndike in which each of the 100 addition combinations taking 1-digit numbers 2 at a time were separately specified as distinct intended learning outcomes: can add 0 + 0, can add 0 + 1, can add 0 + 2, and so on (Fishbein & Tyler, 1973). Tyler defined intended learning outcomes far more generally, at a much higher level, in a manner that would not have been amenable to behaviourist learning approaches. The NVQ model also recommended far more general outcomes. To put it simply, the NCVQ did not base any of its work on behaviourist learning theory, nor was the work of Watson, Guthrie, Thorndike, and Skinner represented at all in the NVQ model (see also Burke, 1995).

Having said that, it is important to appreciate that the radical critique of the NVQ model followed in the wake of a radical critique of the Objectives Movement, which

included trying to argue that even Tyler was a behaviourist.⁹¹ In other words, when England began experimenting with outcome-based qualification design in the 1970s and 1980s, there was already plenty of ammunition available to critics, arising from a vast amount of experience of the Objectives Movement in North America. Some of these criticisms were entirely legitimate, while others were clearly not.

Although this scholarly debate might seem a little arcane, it is important to engage with it explicitly. This is because many of the early criticisms of the NVQ model argued that outcome-based approaches were either fundamentally flawed (for example, the 'naïve behaviourism' critique from this subsection) or required sacred cows to be slain (for example, the 'rejection of learning programmes' critique from the previous section). If true, then this would render outcome-based models both implausible and unworkable. Indeed, this would challenge the very concept of an outcome-based model, which extends the critique of NVQs to all CASLO qualifications, as a matter of principle. However, we remain unconvinced that outcome-based models can be dismissed quite so easily.

Confusing model

Although we have argued that scholars sometimes mischaracterised the NVQ model, it was undoubtedly a confusing model, and its proponents did themselves no favours by failing to resolve fundamental ambiguities before rolling the model out.

The outcomes at the heart of the NVQ model were supposed to be explicated on a principled basis, using functional analysis, which drew upon the Job Competence Model. Unfortunately, the status of underpinning knowledge and understanding proved to be highly controversial, and this ambiguity was never entirely resolved.

According to NVQ theoreticians, it was incorrect to specify underpinning knowledge and understanding as part of the occupational competence construct. Nor, they argued, was it necessary to provide evidence of underpinning knowledge and understanding independently of competent performance. In other words, if that knowledge and understanding genuinely underpinned occupational competence, then evidence of competent performance would, by definition, imply possession of the underpinning knowledge and understanding. The logic seems undeniable.

Yet, when NVQ developers grappled with the pragmatics of assessing competent performance – both authentically and comprehensively – this logic began to unravel. The assessment of knowledge and understanding was soon introduced to the NVQ approach as a substitute for assessing competent performance, either when it was

⁹¹ This is a curious literature, which is illuminated by an exchange between Hlebowitsh (1992), Kliebard (1995) and Hlebowitsh (1995), concerning an influential critique of the Tyler Rationale by Kliebard during the 1970s.

impossible to assess competent performance directly, or as a warrant for generalising competent performance from a single context to multiple contexts. Thus, knowledge and understanding came to play a more central role in the model, albeit even then not as a component of the occupational standard itself. Eventually, knowledge and understanding requirements became a mandatory component of occupational standards. Ultimately, they even came to be assessed via discrete qualifications, known as Technical Certificates.

Even today, it is still not obvious how knowledge and understanding requirements ought to be specified and assessed, although – given the complexity of inferring underpinning knowledge and understanding requirements from an occupational competence model – it would make sense for this not to be left entirely to teachers and trainers to work out. If so, then the burden must fall either to standards developers or to qualification developers. What remains particularly unclear is the extent to which underpinning knowledge and understanding need to be assessed directly, whether to provide indirect evidence of occupational competence, or simply for their own sake. Assessing knowledge and understanding separately, for their own sake, runs its own risks, particularly where knowledge and understanding requirements are separately certificated. Winch, for example, has noted the subtle absurdity of it being possible (by around 2010) for an apprentice to achieve an NVQ before being awarded the certificate that confirmed their acquisition of the underpinning knowledge and understanding required for that NVQ (Winch, 2011).

Shambolic rollout

Implementation of the NVQ model was shambolic. Even accounts provided by key protagonists, including Gilbert Jessup and Graham Debling, give the impression that the NCVQ and the Training Agency were to some extent making the model up on the fly. It is clear that all concerned were very aware of the size of the practical challenges that would need to be overcome (see Burke, 1989, for examples). Yet, the implementation timetable was extremely ambitious.

Implementing functional analysis proved to be especially challenging. Mansfield & Mitchell described it as “probably the most misinterpreted, misunderstood and most haphazardly applied method ever to emerge from the discipline of occupational analysis” (1996, page 98). Some years earlier, Mitchell had warned that “functional analysis is not a method (for the moment at least) which can be taken off the shelf by those with a little time and a handbook and applied well [...] it is an expert system which requires [a] good deal of background understanding” (1989, page 59). Debling, in the same book, had warned that “there is very little expertise in defining explicit national standards” (1989, page 83). Raggatt & Williams summed it up like this:

The attractiveness of the approach notwithstanding, how realistic was it to expect individual lead bodies, even with the assistance of consultants, to be capable of implementing a full-scale functional analysis of the occupations within their respective sectors – or even understanding what it entailed?

(Raggatt & Williams, 1999, page 97)

Raggatt & Williams argued that officials underestimated the complexity and messiness of employment situations and structures (Raggatt & Williams, 1999). Where officials intended a rigid rollout of their highly technical model, employers pushed back. Perhaps inevitably within this largely voluntaristic context – one that was designed to be led by employers – the rollout proved to be very far from rigid. The breadth of the competence model often proved to be a casualty in this struggle between officials and employers, with powerful interest groups exerting a narrowing influence on standards (Callender, 1992).

The shambolic rollout of a confusing model provided much fuel for the fire of critics. Yet, there was still plenty of support for NVQs and the idea of competence standards more generally (FEFC, 1997). Although many employers did reject the NVQ model because of its rigidity, others argued the case for sufficient flexibility to make it work in their own particular circumstances, including the Association of Accounting Technicians (Purcell, 2001). And, while many scholars rejected the NVQ model outright, others felt that it would be wrong simply to abandon NVQs and start again (Hodkinson & Issitt, 1995). Ultimately, the NVQ train kept on rolling.

Rigour-ish

Interestingly, NVQs managed to resist the suggestion that they ought to locate external written exams front and centre. Government asked the Beaumont review to explore options for embracing externality, but no such recommendations followed. Statutory Regulations for the new NQF (QCA, 2000) introduced a requirement for independent assessment, which might include external exams. Yet, a subsequent QCA evaluation indicated the strength of feeling against exams within TCs, and a subsequent SSDA-commissioned evaluation indicated even stronger resistance to external testing of key skills. As TCs became integrated within the QCF, they would have been required to adopt the CASLO approach (as we will see later on). Even by the final edition of the NVQ Code of Practice, the model was still firmly grounded in internal assessment (QCA, 2006). That is not to say that NVQs never included external exams (more frequently known as tests). They sometimes did. Indeed, the original guidance clearly stated that a wide range of techniques could be accommodated, including paper-based questioning, computer-based questioning, and so on (TA, 1989b). But external testing was never central to the NVQ model, as it came to be for other CASLO qualifications, including GNVQs.

The call for more external testing is often associated with concern for greater rigour. Rigour was always a matter of concern for NVQs, which was recognised by Beaumont and echoed by Dearing. Yet, concerns for greater responsiveness – to employers, in particular – tugged in a different direction. For instance, it was clearly not straightforward to rollout a system based on functional analysis. It required considerable technical expertise to apply the technique when developing standards and qualifications. And these standards and qualifications – with their unique logic and grammar – also required a certain amount of technical expertise to facilitate effective use. Where developers saw necessary precision, employers saw unintelligible jargon. Toward the end of the 1990s, it seemed that the employers might have won out, resulting in a dilution of the approach. By 2000, QCA had eliminated the expectation that standards and qualifications should derive from functional analysis. This introduced a new elephant to the room: if not functional analysis, then what? A decade later, though, functional analysis was firmly back on the table, as a requirement of the UKCES quality criteria for NOS (although the UKCES version of functional analysis was not quite as stringent as the version pioneered during the early years).

Responsiveness was always going to be a major challenge for a national qualification that was intended to displace a so-called ‘jungle’ of disparate, bespoke ones. NVQs were designed to be flexible, but neither they nor the apprenticeship frameworks within which they came to be located were seen, by many employers, as flexible enough. In an inherently voluntaristic system where employer engagement was desperately required yet often not forthcoming, it is easy to see how demands for responsiveness may have weighed more heavily than concerns over rigour.

Conclusion

We can safely conclude that some of problems that beset NVQs were due to the particular version of the CASLO approach that was adopted. Although it was quite neat, in theory, the exclusive focus upon competent performance failed to persuade stakeholders. The status of underpinning knowledge and understanding within the NVQ model continued to cause problems for decades. The focus on competence, and the concomitant emphasis on workplace assessment, was meant to provide a welcome antidote to an historical overreliance upon theory and written exams. Ultimately, though, it swung the pendulum too far in the opposite direction, leaving a ‘hole’ that needed to be ‘patched up’ by successive incarnations of knowledge and understanding requirements. It is important to stress that the decision to frame outcomes in terms of occupational competence alone – leaving knowledge and understanding requirements implicit – was unique to the early NVQ model. It is not a feature of outcome-based qualification models, per se, and it was not a feature of many subsequent CASLO qualifications in England.

We can also safely conclude that many of the problems that beset NVQs were due to rushed and poor implementation. In retrospect, the magnitude of the task that faced the NCVQ and associated agencies was mind-boggling. It is hard to imagine how a task of that complexity, in a context as messy as the one in which it was located, could ever have been implemented effectively in the space of just a few years.

What is unclear is the extent to which the problems that beset NVQs were due to adopting the CASLO approach, per se, which would be to question the viability of the approach itself in this context. It may have been a principled decision to base the NVQ model on the CASLO approach, but whether it was an optimal decision under the circumstances is not possible to judge on the basis of historical evidence alone.

GNVQs

The introduction of General National Vocational Qualifications can be understood as a pragmatic response to resistance to the proposal that all vocational and technical qualifications would need to be accredited to the NVQ framework (Sharp, 1998). As such, GNVQs were partly a response to concerns from stakeholders such as the Confederation of British Industry that NVQs were specified too narrowly. But they were also partly a response to concerns from stakeholders such as the Business and Technician Education Council that existing college-based vocational qualifications fulfilled a critical 'middle way' function – between general education and technical training – that NVQs were incapable of serving. By accepting the need for both NVQs and general NVQs, the National Council for Vocational Qualifications formally acknowledged the importance of this middle way.

In May 1991, a few years after the introduction of NVQs, the white paper 'Education and Training for the 21st Century' announced the introduction of General NVQs, high-quality vocational alternatives that would cater for the increasing numbers of young people who were staying in full-time education.⁹² At Level 3, the system would

⁹² Thus, GNVQs followed in the wake of pioneering programmes – such as the Certificate of Pre-Vocational Education and the Technical and Vocational Education Initiative – which had attempted to target 'less academic' young people who wanted to remain in education but who were not particularly well suited to A levels (Sharp, 1998). These qualifications, in turn, were influenced by an early report from the Further Education Unit (see Pring, 1995), which had been established in 1977 by the Secretary of State for Education and Science to facilitate a co-ordinated and cohesive approach to curriculum development in further education. Two years later, the FEU published a report on pre-employment courses for young people entering further education at 16. It was titled 'A Basis for Choice' because it aimed to develop a design template for pre-employment courses that would help young people to make an informed and realistic career choice (FEU, 1979). This would be facilitated by courses based upon a common core of learning – guaranteeing competence in basic skills –

comprise 3 distinct routes – the A level route, the GNVQ route, and the NVQ route – as well as providing opportunities to combine qualifications across routes.⁹³ The white paper explained its reasoning as follows:

Many young people want to keep their career options open. They want to study for vocational qualifications which prepare them for a range of related occupations but do not limit their choices too early. Some want to keep open the possibility of moving on to higher education. Employers, too, want to have the opportunity of developing their young recruits' general skills, as well as their specific working skills. A range of general qualifications is needed within the NVQ framework to meet these needs. Some already exist which help to meet this need – including some offered by the Business & Technician Education Council (BTEC). But they need to be clearly related to the NVQ framework, to make it easier for people to progress quickly to occupationally specific qualifications.

(DES, DE, & WO, 1991, page 18)

Although this seemed to suggest the possibility of accrediting BTECs (and other vocational qualifications) as General NVQs, this proved not to be the case. As with NVQs, GNVQs ended up being developed from scratch. Indeed, under the auspices of the NCVQ, they were designed to be far more like NVQs than A levels. GNVQs were introduced at 3 levels:

- Advanced – comprising 8 mandatory vocational units, 4 optional vocational units, and 3 mandatory core skills units (at Level 3) – which was usually studied as a 2-year programme
- Intermediate – comprising 4 mandatory vocational units, 2 optional vocational units, 3 mandatory core skills units (at Level 2) – which was usually studied as a 1-year programme

alongside vocational and job-specific studies. The report specified aims for this common core in terms of learning outcomes and learning experiences. Linked to a profile approach to reporting, this pre-employment course design template can also be seen as a precursor to the CASLO approach.

⁹³ In his foreword to the white paper, Prime Minister John Major also announced the creation of an overarching Level 3 diploma: “With the introduction of a new Advanced Diploma, we will end the artificial divide between academic and vocational qualifications, so that young people can pursue the kind of education that best suits their needs. While A levels will remain the benchmark of academic excellence, we will raise the standard of vocational qualifications.” (DES, DE, & WO, 1991). Jessup (1993) explained that the purposes of the diploma were to signal parity of esteem for the vocational route and to broaden the post-16 curriculum more generally. Ultimately, the idea of an overarching diploma was abandoned.

- Foundation – comprising 3 mandatory vocational units, 3 optional vocational units, 3 mandatory core skills units (at Level 1) – which was usually studied as a 1-year programme⁹⁴

As students were required to pass all units to achieve their GNVQ – mandatory, optional, and core skills alike – this meant that the GNVQ was effectively a grouped award (unlike A levels, for instance, which had delivered qualifications on a subject-by-subject basis since the 1950s).

Pilots for 5 Advanced and Intermediate subjects commenced in September 1992.⁹⁵ These qualifications were formally launched in September 1993, with Foundation GNVQs following a year later. Ultimately, GNVQs were made available in 14 subject areas by 3 awarding organisations – BTEC (later Edexcel), RSA (later OCR), and City & Guilds (later AQA, with whom City & Guilds had formed an alliance).

According to statistics provided by the Further Education Funding Council for the 1995 to 1996 academic year, the relative split of GNVQ students across these 3 awarding organisations was: 74% (BTEC), 10% (RSA), and 16% (C&G).⁹⁶

The Further Education Unit (FEU) prepared a comprehensive manual on how to implement GNVQs, which emphasised just how significant a change they represented, particularly for teachers who came from the A level route. Borrowing certain ideas from the BTEC model, GNVQs embodied a radical student-centred philosophy. The report characterised the ‘spirit’ of the GNVQ as follows (FEU, 1994, page 198):

- learners are responsible for producing and presenting evidence to show that they have met the performance criteria
- an approach to learning and assessment which is based on the application of skills, knowledge and understanding within ‘holistic’ learning experiences
- the concept of not-yet-achieved, rather than failed

⁹⁴ Advanced GNVQs – more specifically, their 12 vocational units – were intended to be equivalent to 2 A levels (or 1 Level 3 NVQ). Intermediate GNVQs were intended to be equivalent to 4 or 5 higher-grade GCSEs (or 1 Level 2 NVQ). Foundation GNVQs were intended to be equivalent to 4 lower-grade GCSEs (or 1 Level 1 NVQ). In September 1995, Part One GNVQs were introduced (not to be confused with Part One NVQs). They were designed to provide a 2-year course for key stage 4 students or a 1-year to 2-year course for post-16 students. They required less teaching time than Foundation and Intermediate GNVQs, but shared many characteristics in common (Frankland & Ebrahim, 2001).

⁹⁵ Art & Design, Manufacturing, Leisure & Tourism, Business, and Health & Social Care.

⁹⁶ This contrasted with essentially the same data for NVQs: 8% (BTEC), 24% (RSA), and 68% (C&G). Note that data for RSA and C&G related to entries, whereas data for BTEC related to awards (FEFC, 1997).

- the concept of mastery learning
- an emphasis on the assessment of skills, knowledge and understanding through their application

GNVQs were intended to offer considerable scope for personalisation, including via optional units or additional units. In theory, at least, a GNVQ course was not time limited, suggesting that students could learn at their own pace, with the potential to enter and exit at different points in the year (which is how BTECs typically operated). Personalisation was also embodied in the principle that students should begin with an initial diagnostic assessment of core skills needs, resulting in an individual profile of strengths and weaknesses and an individualised action plan. This diagnostic assessment was also intended to explore the potential for Recognition of Prior Learning in relation to the broader GNVQ programme. Reflecting on the scale of change anticipated, the report observed that: “GNVQs may require a significant culture shift, as well as the development of new staff skills” (FEU, 1994, page 2).

Design

The GNVQ design process was led by Gilbert Jessup, who proposed essentially the same outcome-based approach as had been pioneered within NVQs. NCVQ Chief Executive, John Hillier, once claimed that the GNVQ model represented the: “most extensive application of outcomes-based assessment in the world” (Ecclestone, 2002, page 3). GNVQs were based almost entirely on the CASLO approach.

Standards for all units – both vocational and core skills – were specified via elements (learning outcomes) and performance criteria (assessment criteria), which were related to National Occupational Standards in relevant sectors (Hodgson & Spours, 2003). Figure 8 illustrates an element, associated performance criteria, a range statement, and evidence indicators from the original specification of an Advanced vocational unit in Business (from Allen, 2004, Appendix 1). Figure 9 illustrates an element and associated performance criteria from the specification of a Level 2 communication core skills unit (from FEU, 1994, page 178).

Each element also had a statement of range, which indicated the significant dimensions that had to be covered and evidenced. The kinds of evidence required for each element were described via evidence indicators. According to the GNVQ implementation manual:

Assessors need to develop the ability to apply all the relevant criteria to a piece of candidate’s work as a global judgement, rather than as a process of ticking off disaggregated criteria.

(FEU, 1994, page 209)

Unit 1 Business in the economy

Element

1.1 Explain the purposes and products of business

Performance criteria

- 1 Demand for goods and services is identified and described
- 2 Demand in relation to particular product is described
- 3 Industrial sectors are identified and described
- 4 The product of businesses in different industrial sectors is identified and described
- 5 Purposes of selected business organisations are explained

Range

Demand: needs, wants and effective demand, consumption and income, demand and price, elastic and inelastic

Industrial sectors: primary, secondary, tertiary

Product: goods, services

Purposes: profit-making, public service, charitable

Evidence indicators

An analysis of selected businesses with an explanation of why businesses exist, an explanation of their product and an explanation of demand in general and demand in relation to a particular product. Evidence should demonstrate understanding of the implications of the range dimensions in relation to the element. The unit test will confirm the candidate's coverage of range.

Figure 8. Advanced vocational unit in business

- 2.2 Prepare written material on routine matters
- PC 1 All necessary information is included and information is accurate
 - PC 2 Documents are legible.
 - PC 3 Spelling, grammar and punctuation are used correctly.
 - PC 4 The format used to present material is appropriate and information is ordered appropriately to maximise audience understanding.

Figure 9. Level 2 communication core skills unit

Aware of earlier criticism of the NVQ approach, the FEU report was alive to the risk of fragmented learning arising, indirectly, from the manner in which criteria were nested within outcomes that were nested within units. It insisted that neither learning nor assessment needed to be similarly disaggregated, arguing instead for activities and assignments that represented coherent and rounded experiences for students. The report distinguished between unit-based and integrated delivery programmes, noting that unit-based ones were easier to deliver but more likely to lead to fragmented learning and assessment. Integrated programmes – which pulled together elements from a number of units – mitigated this risk, although they did make it harder for students to track their achievements.

Assessment typically involved a combination of extended assessment, with evidence accumulated in a portfolio, plus externally set tests.⁹⁷ Students would complete their portfolio largely on the basis of assignments and activities developed by their teachers. For the purpose of internal and external quality assurance – known as verification – students would index their portfolio to relevant elements and criteria. The portfolio would need to demonstrate how each performance criterion had been satisfied for each element (and therefore how all elements had been achieved for each unit) across all vocational and core skills units. Evidence might take a variety of different forms – assignment reports, artefacts, diagrams, videos, witness testimony, and so on.

To pass their GNVQ, a student would need to satisfy all performance criteria for all elements of each vocational and core skills unit. In addition, they were required to pass an external test for each of the mandatory units (although a few mandatory units were excluded from this requirement where content was thought to be inappropriate for testing). These multiple-choice tests were set by awarding organisations and either externally marked or internally marked with verification. They were intended to supplement the portfolio of evidence, assessing underpinning knowledge and understanding. The tests were designed to confirm broad coverage of range.⁹⁸ Technically, they were based on a compensatory principle, as passing the test depended solely on having achieved the overall pass mark. However, they still embodied a loose conception of mastery:

⁹⁷ The NVQ model was not actually based upon continuous assessment, per se, because there was no formal link between the chronology of learning and the chronology of assessment. Indeed, it would have been entirely in keeping with the model for all of the assessment to have occurred after all of the learning had been completed. However, it left open the possibility of a more extended, or continuous, assessment process where certain outcomes were mastered earlier or later than others.

⁹⁸ Wolf (1998) described these simple tests (which played no part in grading) as little more than a concession to ministers, noting that the original NCVQ model for GNVQs was entirely portfolio based.

Because of the concept of 'mastery learning' a high pass mark has been set (currently 70%), and it is envisaged that students will repeat the test until they achieve it. Tests are available several times a year to enable students to take the relevant test when they are judged to be ready to do so.

(FEU, 1994, page 213)

The idea of mastery learning was central to the original GNVQ model. It was embodied in the design principle of fusing formative and summative assessment, which was intended to empower students to maximise their achievements (Ecclestone, 2000).

Although individual units were not graded, there was a process for deriving an overall qualification grade, such that students would receive either a pass, a merit, or a distinction grade overall. Grades were based on the same portfolio of evidence that determined whether or not a student passed, although different criteria were used for awarding merit and distinction grades. Originally, the Advanced GNVQ had 6 grading criteria – that is, a set of 6 for merit and a parallel set of 6 for distinction – grouped within 3 themes. Focused primarily on processes, it was soon recognised that another theme would be required to recognise the overall quality of work produced. So, by September 1994, there were 8 grading criteria, grouped within 4 themes, as presented in Figure 10 (from NCVQ, 1994, page 27). As described in the booklet that specified these new criteria:

The grading criteria focus on students' performance because GNVQs are designed to encourage active approaches to learning; how students tackle their work; how much responsibility they take for planning it; how they decide what information they need; how well they review and evaluate their performance; and the overall quality of the work they produce.

(NCVQ, 1994, page 8)

Recognising that a student would acquire these skills gradually throughout their course, and acknowledging that not all assignments would elicit evidence of higher-level performance, the NCVQ specified that criteria for merit or distinction grades only needed to be demonstrated across one-third of the portfolio of evidence. Thus, students could be awarded a merit grade if one-third or more of their evidence met all of the merit grading criteria, or a distinction grade if one-third or more of their evidence met all of the distinction grading criteria.⁹⁹ Again, the requirement that all criteria needed to be satisfied for the award of a higher grade was consistent with the idea of mastery, albeit operationalised pragmatically rather than absolutely.

⁹⁹ The idea of one-third of the evidence meeting criteria was not intended to be interpreted mechanistically, but heuristically.

Theme	Aspect	Merit criterion	Distinction criterion
Planning	1. Drawing up plans of action	Student independently draws up plans of action for a series of discrete tasks. The plans prioritise the different tasks within the given time period.	Students independently draws up plans of action for complex activities. The plans prioritise the different tasks within the given time period.
Planning	2. Monitoring courses of action	Student independently identifies points at which monitoring is necessary and recognises where revisions to courses of action are necessary. Appropriate revisions to plans are made with guidance from teacher/tutor.	Student independently identifies points at which monitoring is necessary and recognises where revisions to courses of action are necessary. Appropriate revisions to plans are made independently.
Information seeking and information handling	3. Identifying and using sources to obtain information	Student independently identifies, accesses and collects relevant information for a series of discrete tasks. Student identifies principal sources independently and additional sources are identified by the teacher/tutor.	Student independently identifies, accesses and collects relevant information for complex activities. Student uses a range of sources, and justifies their selection.
Information seeking and information handling	4. Establishing the validity of information	Student independently identifies information which requires checking for validity. Student checks validity of information using given methods.	Student independently identifies information which requires checking for validity. Student independently selects and applies appropriate methods for checking validity.
Evaluation	5. Evaluating outcomes and alternatives	Student judges outcomes against original criteria for success; identifies alternative criteria that can be applied in order to judge success of the activities.	Student judges outcomes against original criteria for success and identifies and applies a range of alternative criteria in order to judge success of the activities.

Theme	Aspect	Merit criterion	Distinction criterion
Evaluation	6. Justifying particular approaches to tasks/activities	Student justifies approach used; indicates that alternatives were identified and considered.	Student justifies approach used, basing justification on a detailed consideration of relevant advantages and disadvantages. Alternatives and improvements are identified.
Quality of outcomes	7. Synthesis	Student's work demonstrates an effective synthesis of knowledge, skills and understanding in response to discrete tasks.	Student's work demonstrates an effective synthesis of knowledge, skills and understanding in response to complex activities.
Quality of outcomes	8. Command of language	Student's work demonstrates an effective command of the language of the GNVQ area at Advanced level.	Student's work demonstrates a fluent command of the language of the GNVQ area at Advanced level.

Figure 10. Advanced GNVQ grading criteria

Evolution

The GNVQ model was both novel and complex. The fact that it was designed and implemented at speed resulted in numerous problems, and the model had to be reconfigured several times over a period of just a few years. In his Foreword to the GNVQ manual, FEU Chief Officer, Geoff Stanton, emphasised the “hectic pace” at which GNVQs had been introduced, adding that there was no good time at which to produce guidance because the continual revision of approaches and specifications quickly rendered any such guidance out of date (FEU, 1994, page 1).

It is hard to say exactly how many discrete models were implemented between its conception and demise – because it is hard to say which changes were mere refinements and which were substantive reforms, especially as implementation was staggered across subject areas – but Ecclestone (2002) proposed that there were 4 distinct models, with the following dates corresponding to their first teaching:

1. September 1993 model
2. September 1995 model
3. September 1996 model
4. September 2000 model

The final model involved a fully-fledged reform process, giving birth to a new qualification that was to become known as the Advanced Vocational Certificate of Education (AVCE).

The first model was heavily criticised from the outset, including within the high profile critique of the NVQ-GNVQ system mounted by Alan Smithers, which we considered in the last section (Smithers, 1993). Further Education Funding Council inspectorate surveys subsequently identified concerns over an unwieldy assessment system, inappropriate and unclear external test questions, poor teaching of key skills, inadequate internal and external verification, and some low completion rates (FEFC, 1994b; 1995). Office for Standards in Education inspection reports also identified similar problems (Ofsted, 1993; 1994). Independent evaluations identified further concerns (Wolf, Burgess, Stott, & Veasey, 1994; Wolf, Scharaschkin, Meade, & Pettitt, 1994). These reports differed in their appraisal of how serious the challenges were. The FEFC, for instance, referred to “a number of teething problems” (FEFC, 1994b, page 5), whereas Wolf – focused specifically on grading – outlined the need for “major reconceptualisation and reform and not simply fine-tuning” (Wolf, Burgess, Stott, & Veasey, 1994, page 1).

As early as March 1994, Under-Secretary of State for Further and Higher Education, Tim Boswell, had set out a 6-point agenda for action for the NCVQ, to ensure quality and rigour in GNVQs. This was the beginning of an extended period of refinement,

review, and reform. The 2 most important reviews of this period were the 'GNVQ Assessment Review' (Capey, 1995) and the overlapping 'Review of Qualifications for 16-19 Year Olds' which had a broader remit (Dearing, 1996). The Capey review was particularly significant for the future of GNVQ. It contained both immediate and longer term recommendations. The longer term ones, once piloted, resulted in GNVQ model 4 (the AVCE).¹⁰⁰ Its immediate recommendations influenced the design of GNVQ model 3.

The report of the Capey review strongly supported a number of GNVQ design features, including the specification of learning outcomes, the unit-based structure, the emphasis on active learning, and the inclusion of core skills. However, it also took issue with various aspects of the assessment regime, as this was configured in the September 1995 model. Changes for September 1995 had already included reducing assessment documentation to a minimum, with recording at the element level not at the performance criterion level, and no longer requiring recording of range coverage. The report identified 4 concerns that were common to most submissions to the review:

- the (continuing) very serious burden of assessment and recording on teachers and students
- interpretation and application problems associated with the grading criteria
- difficulties in teaching and assessing core skills
- uneven quality and demand of the external tests

The problem of burden was related to a heavy continuous assessment load, associated with the CASLO approach. A survey by the Association for Colleges suggested that, while there was support for criteria-based assessment, there was also concern that too much time was spent assessing, which took time away from teaching and learning.

Of particular relevance to the CASLO approach, Capey recommended that:

1. GNVQ assessment should (in the longer term) move from element-based to unit-based assessment
2. NCVQ should investigate the feasibility of external tasks for core skills units
3. the purpose of the tests should be reviewed, along with the 70% pass mark, and whether they should contribute to grading

A key consideration addressed by the Review was the role of mastery in a general qualification such as the GNVQ. For instance:

¹⁰⁰ The pilot of the fully revised GNVQ model began in September 1997 and reported 2 years later (FEFC & Ofsted, 1999).

The fundamental issue was whether the GNVQ model could move away from the conventional mastery model to one which would identify and assess the key knowledge and skills within a unit. The attendant problems of any such moves could be to reduce the transparency of the required outcomes for both students and users (ie what will the student have covered from the range?) The advantages are that the unproductive work of superficially covering all the range could give way to more in-depth work on a more focused range of skills and knowledge.

The group agreed that the move towards sampling rather than exhaustive coverage, which began in the September 1995 changes, should continue. However, this would not resolve the problem of the large number of evidence indicators prescribed for each unit (up to 50 in some cases). This could be eased by assessing at **unit**, rather than **element** level, leaving the content unchanged but reducing the number of evidence indicators by adopting a more integrated approach. The group thought that this might also lead to more effective learning. One implication is that assessments would not necessarily cover all aspects of the elements, performance criteria and, particularly, range. Another implication is that assessors would be making more generalised judgements about performance (see the level descriptions in National Curriculum assessment) when unit-referenced evidence indicators replaced element-referenced indicators.

(Capey, 1995, page 24)

The subsequent report of the Dearing review was less detailed but more blunt:

A qualification like the NVQ should not be granted unless a candidate has demonstrated all the competences necessary to provide a reliable service to a client. But the GNVQ is not a professional qualification: it covers a broad area of knowledge and understanding which underpins a range of trades and professions, and provides a basis for a practical education. The mastery model is therefore inappropriate to the GNVQ, so I welcome the proposal in the Capey Report that assessment should no longer cover every detail, but it should be based on an overall assessment of performance in defined areas known as units. This will reduce workload and avoid the risk that assessment may become a burdensome series of atomistic assessments.

(Dearing, 1996, page 77)

It is interesting to note how Dearing, in particular, dismissed (or simply overlooked) the potential significance of mastery learning in general education, as though the concept of mastery was relevant only to occupational qualifications as a certification requirement. We will return to this point later.

Completion rates were problematic from the outset, and continued to be so from one model to the next. For instance, although the majority of full-time students in further

education colleges who completed GNVQ courses passed them, pass rates were not particularly high. For the 1995 to 1996 academic year, these ranged from 62% for Foundation GNVQs to 74% for Advanced GNVQs (FEFC, 1997). This FEFC report also raised concern over large numbers of students dropping out before even taking their final assessments. Commenting on this issue, Wolf (1998) argued that an important factor in explaining non-completion was the failure of many students to keep up a steady rate of portfolio completion. Echoing an earlier evaluation, she concluded that the GNVQ approach relied considerably on the ability of tutors to organise and manage student learning.

Standards remained an issue for GNVQ delivery, as discussed in some detail by the 1997 FEFC report, as well as by a 1998 Ofsted report on Advanced GNVQs (relating to the 1996 to 1997 academic year). While both the FEFC and Ofsted identified plenty of good practice, they both expressed concerns related to how consistently standards were being applied. The Ofsted report concluded that most students had enjoyed their courses, estimating that the majority had achieved more in GNVQ than they might reasonably have been expected to achieve at A level. However, the report also expressed continuing concern over the grading system and verification procedures, concluding that the:

most serious weakness which must be addressed is that the lack of clearly defined standards results in over-generous assessment and grading by some teachers and verifiers.

(Ofsted, 1998, page 5)

It added that external verification by non-subject specialists contributed to this unreliability.

Ofsted summarised what seemed to be a promising situation just prior to the introduction of model 4 in its 1999 Annual Report (Ofsted, 1999). GNVQs had improved, the new assessment regime was more rigorous, and standards of performance had improved accordingly. Students on Foundation and Intermediate courses were particularly motivated by the links with vocational sectors, although achievement was more consistent on Intermediate courses, where the highest attaining students benefitted from the independent style of learning promoted by the course. Good planning was judged to be an essential component of effective provision – where planning was poor, pupils frequently failed to complete their portfolio work. In school sixth-forms, the large majority of Advanced GNVQ students were committed and conscientious, with frequent contact with the world of work promoting a high level of self-reliance and interpersonal skills. The model 4 pilot had gone well: the new assessment approach was easier to manage and appeared to be encouraging greater rigour, leading to improved performance standards.

AVCEs

Following this period of piloting, the model 4 Advanced GNVQ was rolled out in September 2000, now with a new name, the Advanced Vocational Certificate of Education (AVCE). According to the QCA (2003), key changes involved separating out key skills, so that candidates no longer failed the GNVQ if they failed to achieve the key skills, and introducing a compensatory approach to assessment, such that candidates were no longer required to pass every unit to achieve the qualification.

The AVCE was introduced in tandem with new A level qualifications, as a key part of the Curriculum 2000 reform programme, which arose from the 1997 Department for Education and Employment 'Qualifying for Success' consultation. A key intention underlying this reform programme was to address undue narrowness and lack of flexibility in the post-16 curriculum (QCA, 1999c). Students would be encouraged to study for both GNVQ and A level within the same Level 3 programme, facilitated by comparably sized units and the new GNVQ grading scale (A to E).

The AVCE was made available in 3 sizes: 12 units, 6 units, and 3 units. The 12-unit AVCE, known as the double award, was intended to be of a standard equivalent to 2 A levels, comprising a minimum of 6 and a maximum of 8 compulsory units, with a maximum of 6 optional units (Ofsted, 2004).

According to the QCA (1999b), the new assessment arrangements were designed to increase rigour and manageability. One-third of the assessment would typically be external (involving set assignments or tests), while two-thirds would typically be internal with moderation (based upon a portfolio of evidence).

In relation to the CASLO approach, it is important to emphasise that AVCE units no longer specified learning outcomes as elements. Like their A level counterparts, unit content was specified primarily through syllabus content. In short, the new AVCE-GNVQ was no longer outcome-based.

Having said that, there were still echoes of the CASLO approach in the specification of unit standards via grading criteria. Grading was no longer a simple process of ticking off 'the bullets' as per previous models (see Ecclestone, 2002). Instead, points were available for performances within each unit, and these points were aggregated across units to form an overall point total, from which the qualification grade was derived. This also removed the formal requirement for students to pass all AVCE units for the award of an overall qualification grade, just as long as their

overall point total exceeded the grade E threshold. In short, the new AVCE-GNVQ was no longer mastery-based.¹⁰¹

The process for deriving points for each centre-assessed unit was complex, and did actually retain an element of mastery. Criteria were specified for unit grades E, C, and A, and whether or not these criteria had (all) been satisfied determined the range of unit points available to students. For example, the following criteria were specified for an early Advanced business unit:

E1 classify the business according to its product or service

E2 describe and explain the objectives of the business

E3 describe the functional areas that exist in the business, and explain how they help the business to meet its objectives

E4 describe the management styles and cultures present within the business

E5 identify communication channels used by the business

E6 explain how the production process and quality assurance/control system used by the business helps it to add value to its product or service

(Allen, 2004, Appendix 2)

If a student failed to satisfy all of these criteria, then their unit points were limited to 0 to 6, depending on how many of the grade E criteria had been satisfied. Conversely, if the student had achieved all of these criteria, then additional points were available to them. If they had achieved all of the grade E criteria, but had not achieved all of a separate set of grade C criteria, then points 7 to 12 would be available to them, depending on how many of the grade C criteria had been satisfied. For a full description of this process, see Greatorex (2001). The main point to note is the echo of the CASLO approach, as captured in the following evaluation conclusion:

¹⁰¹ Although not discussed in detail here, it is worth noting that Foundation and Intermediate GNVQs were not reformed on the same timescale, which meant that consultation exercises could draw upon experiences of AVCE implementation. In a report on exploratory work of this sort, the QCA noted: "There was a large body of opinion in favour of replacement qualifications having a compensatory assessment structure. It was regarded as a positive development within the 2000 GNVQ specifications. Many consultees were most concerned about having assessments that were fit for purpose as well as realistic and achievable. Some providers were concerned that too much teacher-led assessment would be unmanageable and that an appropriate balance between teacher assessment and external assessment would need to be achieved. A unitised approach was valued." (QCA, 2003, page 12). Note that this work was undertaken in the context of a decision, in July 2000, by the Secretary of State for Education, to replace Part One, Foundation and Intermediate GNVQs with 'GCSEs in vocational subjects'. Following concerns that these might not cater adequately for post-16 students, the decision to withdraw GNVQs was delayed until appropriate replacement qualifications could be identified (QCA, 2003).

On the positive side, a minority thought that the specifications were clearer and that the external assessment was fairer and more rigorous than the type of assessment in GNVQ. On the negative side, the majority of teachers found that the AVCE combined the worst of both worlds – external tests that demotivated learners and put them under pressure, combined with echoes of an NVQ assessment methodology which insisted on coverage of grade-related criteria and extensive portfolio evidence.

(Hodgson & Spours, 2003, page 112)

Sharp characterised the transition from model 1 to model 4 as “a continual series of attempts to escape from the constraints of the original” model, moving gradually toward a “more conventional approach to content, knowledge, assessment and curriculum structure” (Sharp, 1998, page 309). This gave him hope for the long-term future of the GNVQ. Ultimately, though, the AVCE model failed. A damning Ofsted report concluded that:

The AVCE is not well designed. It is neither seriously vocational, nor consistently advanced. The aims of the AVCE are not clearly understood by many teachers and students. We observed a good deal of work that was trivial, as well as some that was excessively demanding.

(Ofsted, 2004, page 5)

Even after almost all of the remnants of the CASLO approach had been removed, problems persisted:

Teachers were particularly constrained by the AVCE assessment requirements. Both they and students regard the assessment regime as excessively complex, bureaucratic and hard to understand. They are right. The Qualifications and Curriculum Authority (QCA) has on several occasions attempted to address this issue, but teachers still spend too much of their time assessing, rather than teaching, students. For their part, students spend too much time completing assessments rather than learning.

(Ofsted, 2004, page 5)

AVCEs were quickly replaced by Applied A levels, which were introduced for first teaching in September 2005. As the name suggests, they were even more closely allied to A levels than AVCEs. Put simply, they were A levels. Subsequently, an entirely new vocational qualification, the Diploma, was introduced in September 2008. Like the original GNVQ, this was effectively a grouped award, which required successful completion of all of the components within the programme. Unlike the GNVQ, the CASLO approach was not a significant feature of vocational units within

the Diploma programme.¹⁰² A relatively small suite of Applied A levels continued alongside the Diploma.

Conclusion

Just as for NVQs, GNVQ rollout was highly problematic. This is perhaps not surprising, given that they were part of the same reform programme, and particularly given that GNVQs were something of an afterthought in this process. GNVQs transitioned through a succession of models until they were eventually withdrawn. But, to what extent can we lay the blame for their ultimate demise at the feet of the CASLO approach?

Well, there is certainly some truth in Sharp's observation that successive models represented a retreat from the CASLO approach to a more classical one. Under the CASLO approach, the assessment system was often felt to be unwieldy, if not unmanageable, and problems of inconsistently applied standards were frequently identified. This was particularly true of early implementation, leading Wolf, Burgess, Stott, & Veasey (1994) to some very extreme conclusions, such as:

The unmanageability of the system is not simply a transitional phenomenon but derives from the nature of the assessment process.

(Wolf, Burgess, Stott, & Veasey, 1994, page 2)

The current system is so complicated that it is virtually impossible for all the players in the system – NCVQ, Awarding Bodies, their external verifiers and centres' internal verifiers – to reach consistent conclusions and interpretations.

(Wolf, Burgess, Stott, & Veasey, 1994, page 2)

Overall, we conclude that current grading procedures are generating very different judgement and grading levels between project centres in at least some vocational areas.

(Wolf, Burgess, Stott, & Veasey, 1994, page 5)

On the other hand, it is also fair to conclude that some of the earliest problems would certainly have been due to rushed implementation, combined with insufficient piloting, and insufficient attention to upskilling centres and other participants.

¹⁰² The overall Diploma grade was determined on the basis of performance across Principal Learning components and the Project (Ofqual, 2008c). This involved aggregating points derived from marks, that is, each component involved numerical marking, which proscribed the direct grading of components. So, although there was a nod to the mastery principle in the requirement for candidates to pass all components of the Diploma, the CASLO approach did not feature in this model.

In addition, as for NVQ rollout, it seems safe to conclude that some of the problems that beset GNVQs were due to the particular version of the CASLO approach that was adopted. The complex CASLO-based grading scheme would certainly fall into this category.

Finally, it is important to stress that – despite all of the challenges associated with the rollout of GNVQs across multiple models – the qualification was reasonably popular with both teachers and students. Inspection reports consistently concluded that students who adjusted to the GNVQ approach actually performed very well, which included Advanced GNVQ students performing at a level comparable with their A level peers. Furthermore, it was not until almost all of the remnants of the CASLO approach had been eliminated from the model that the qualification ultimately failed. Its assessment regime was still judged to be excessively complex, bureaucratic, and hard to understand, yet its purpose was no longer clear, and its popularity had declined. This certainly points to wider problems for which it would be unfair to blame the CASLO approach.¹⁰³

BTECs

The Haslegrave report anticipated that a single national council would be established at some point in the future, and this happened in 1983 when the TEC and the BEC merged to form the Business and Technician Education Council (which we shall refer to as ‘the Council’ in subsequent sections).¹⁰⁴ BTECs were to become the largest and for many the most familiar brand of CASLO qualification.¹⁰⁵

¹⁰³ It is fair to say that, while the foregoing account has explained how the GNVQ-AVCE model changed over time, its account of why the model ultimately failed is limited. Responding to an early draft of this report, both Tina Isaacs and Barry Smith emphasised 2 key issues: debate over the nature of Part One (key stage 4) GNVQs, and the merger between the NCVQ and SCAA. Both of these were associated with classical views on qualification design becoming more forceful, resulting in the GNVQ model being bent into the shape of a traditional school-based qualification (see also Oates, 2010). In the case of Part One GNVQs, this ‘academic drift’ was linked to the need to comply with GCSE statutory orders, and to the inevitable challenge of qualifications being taught by school teachers who lacked experience in industry or commerce.

¹⁰⁴ In 1991, this name was changed to the Business and Technology Education Council, at which point it was still a non-departmental public body. The BTEC subsequently became wholly independent of government in October 1993 (Smith, 1994). In 1996, it merged with London Examinations to become Edexcel, which was later acquired by Pearson. The BTEC brand was retained throughout.

¹⁰⁵ It is worth noting that only a few awards (including BTEC and NVQ) have ever entered the public vernacular as shorthand for a certain kind of vocational or technical qualification.

Policies and priorities

In the autumn of 1984, following a public consultation that had been conducted during the spring, the Council published a document that set out its 'Policies and Priorities into the 1990s' (BTEC, 1984). It proposed to consolidate the strengths of TEC and BEC provision and to eliminate any weaknesses.

The Council would continue to exercise its authority by establishing qualifications and qualification standards, by approving centres, and by validating courses – just as the TEC and the BEC had done previously. However, there would be a number of evolutionary changes reflecting developments in industry and commerce, new approaches to pedagogy and assessment, and lessons learnt from monitoring, evaluation, and review. Five principles that underpinned existing TEC and BEC qualifications were formally endorsed, including:

- (a) that one role of a validating body is to promote the important partnership between education, employers and professional interests in designing, developing and reviewing courses and units;
- (b) that the main value of vocational study is demonstrated by what a student can subsequently do and achieve: this requires BTEC to specify the curriculum in terms of the intended outcomes of students' learning;
- (c) that the expectation is that a student who is recruited with integrity to a course should, with diligent study and application, attain a qualification;
- (d) that the assessment of a student is subordinate to, but supportive of, the purpose of a course, so that assessment confirms achievement of learning;
- (e) that the appropriateness of a programme of study lies principally in its relevance as a preparation for success at work, with progression to other studies being important but normally subordinate.

(BTEC, 1984, page 10)

The Council confirmed that units would continue to be designed around "knowledge and skills which the student must attain" (BTEC, 1984, page 13), although it committed to reviewing the way that learning outcomes were articulated.

Without wanting to impose uniformity for its own sake, the Council committed to exploring options for improving synergy across its commercial and industrial qualifications. This included moving towards a single unit and course structure, and adopting a common grading structure across National and Higher National qualifications (which involved grading units but not the overall qualification). Indeed, BTEC had already decided that all higher qualifications would reclaim the Higher

National Certificate (HNC) and Higher National Diploma (HND) nomenclature, which had been associated with earlier (superseded) qualifications (Bourne, 1984).

The following assessment policies are worth highlighting:

82 Assessment is part of the learning process. It should be related both to the aims of the course as a whole and to the objectives of the course's individual components. Assessment confirms the outcome of learning and is the professional responsibility of the teacher.

83 The assessment methods most appropriate to the units and courses are to be used. These may include examinations, tests, vivas, practical work, projects and assignments.

84 The assessment should relate to the student's work throughout the course and should cover all the main elements of study.

85 There should be a sensible balance between intermediate and final assessment, and between formal examinations and other approaches.

(BTEC, 1984, page 19)

The centrality of real world skills development was underlined by BTEC Chairman Neale Raine when he explained that: "the test of vocational education must be what a person can do as a result – not just what can be repeated in a written examination" (Raine, 1984, page 74). He lamented the tradition of forcing learners to acquire knowledge that could only be justified on the basis that it happened to appear in the syllabus of an examination course.

In 1986, the Council published general guidelines on 'Assessment and Grading', 'Teaching and Learning Strategies', and 'Common Skills and Core Themes'. Principles outlined in these documents captured the emerging philosophy of BTEC qualifications, which embraced heavy use of projects and assignments, team working, work-related problem solving, active involvement in own learning, plus stimulating and personalised teaching and learning strategies.¹⁰⁶

Useful insights into these approaches were provided by an NFER evaluation of BTEC Nationals in Business and Engineering that was conducted during 1987 (FEU & BTEC, 1990). The study found much good practice in adapting to this relatively new philosophy, but also many areas for concern. For instance, although the outcome-based approach to qualification design foregrounded competence, applied knowledge, and skill – facilitating an assignment-led approach and discouraging

¹⁰⁶ By 1986, the Council had established a Staff Development Unit to support colleges in coming to grips with these innovative teaching and learning approaches, which emphasised the role of teacher as facilitator and the role of student as explorer (Judith Norrington, personal communication).

overload with non-essential knowledge – there was still a marked tendency for tutors to fall back on more traditional didactic approaches:

‘Getting through the course’ and giving ‘needed’ theoretical knowledge were often viewed as paramount, leaving little time for activity-based work. There was some evidence that theory is often included through habit rather than because it is essential for competence. This seems to be, at least in part, a habit inherited from past course objectives and difficult to break.

(FEU & BTEC, 1990, page 3)

The BTEC qualification suites proved to be very popular. Tables published in ‘BTEC Bulletin No. 3’ (BTEC, 1985) showed that BTEC registrations had increased at all 3 levels – General, National, and Higher National – from 1981 (151,660) to 1983 (181,513).¹⁰⁷ During the 1983 academic year, just over a sixth of registrations were for Generals, just under a third were for Highers, and just over a half were for Nationals. Across the sectors, 40% of registrations were in Business & Finance, 29% were in Engineering, 8% were in Construction, and the remaining sectors accounted for no more than 6% of registrations each. The 16-19 General award was soon to be replaced by the BTEC First. Registrations for Firsts, Nationals, and Higher Nationals continued to rise throughout the 1980s, and by the end of the decade their influence was such that it could be said that they: “formed the heart of the curriculum of most further education colleges” (Higham, Sharp & Yeomans, 1996, page 83).

The RVQ, NVQs, and GNVQs

Just 7 months after the Council had set out its stall on the future of BTEC qualifications, ministers announced a wide-ranging Review of Vocational Qualifications, which would address what was described as a qualification ‘jungle’ (currently presided over by a large number of examining and validating bodies, professional bodies, and other standards-setting bodies, including the BTEC). The scope of this review would range from qualifications that targeted 16-year-old school leavers to qualifications pitched at Higher National level.

Although, in theory, the review was intended to span all vocational qualifications in this range, the inclusion of BTECs was not uncontroversial. This appeared to reflect tension between the Department of Employment (which was driving the review) and the Department of Education and Science (which oversaw the Council). The chair of the review, Oscar De Ville, later suggested that the DES had influenced the Council to steer clear of this employment-driven initiative (Hargraves, 2000). Bear in mind that the TEC and the BEC had been established specifically to co-ordinate and

¹⁰⁷ These figures included both Certificate and Diploma registrations at all 3 levels.

sustain subsequent reforms on a national basis, so the very idea of an independent Review of Vocational Qualifications might have seemed like a vote of no confidence in the newly merged Council (Raggatt & Williams, 1999).

Once responsibility for developing the National Vocational Qualification framework had passed to the NCVQ, it became increasingly clear that existing qualifications would have to be radically reformed if they were to be accredited. Although, in theory, the NVQ model was not wedded to a specific delivery approach, in practice, it was clearly aligned to workplace training and assessment (as envisaged by Gilbert Jessup). As BTECs were normally provided in colleges rather than workplaces, this immediately put them under pressure. Was the NVQ framework designed to exclude or occlude the college-based space (midway between academic A levels and occupational apprenticeships) that BTEC Nationals had so effectively occupied in recent years? It was unclear.

In 1989, BTEC chief executive, John Sellars, was openly critical of early NVQ implementation efforts, which he claimed had resulted in narrow and mechanistic qualifications (Sharp, 1998). The Council was not entirely persuaded by the reform of vocational qualifications, despite other awarding organisations, including City & Guilds and the RSA, having largely bought into it. Nevertheless, the NCVQ and the BTEC published a joint statement in September 1988 explaining how they would work together. The Council then began submitting BTEC Firsts, Nationals, and Higher National awards in Business and Finance and Public Administration for conditional accreditation, explaining that:

Although the representation of course structures will obviously be affected by the project it is not expected to change the major principles and methods by which courses leading to BTEC awards are delivered.

(BTEC, 1989, page 2)

Unfortunately, these early awards were not reaccredited and relations between the Council and the NCVQ remained strained. Under political pressure, a formal agreement was reached in October 1990 whereby the Council agreed to revise its awards to satisfy NVQ accreditation criteria, while still retaining broader, more educational content characteristic of BTECs (Raggatt & Williams, 1999).

As it turned out, the middle route was not eliminated by the NVQ framework, but reinforced, with the introduction of General National Vocational Qualifications. Yet, this established a different kind of threat. GNVQ specifications were to be drawn up by the NCVQ, and it was expected that the BTEC, City & Guilds, and the RSA would award them (Sharp, 1998). Although the expediency of conditional accreditation had given the impression that the BTEC model might continue under the new framework – and although early conversations around ‘general’ NVQs seemed to have left space for integrating BTEC Nationals within the framework (Raggatt & Williams,

1999) – it soon became clear that the BTEC model would have to be replaced by either the NVQ model or the GNVQ model. Indeed, the Council was required to report to the NCVQ annually on its progress in removing BTEC awards (Sharp, 1998).

The NCVQ had no formal powers to require awarding organisations to replace their existing qualifications with new ones that satisfied their accreditation criteria. But the Secretary of State for Education was able to exert pressure via funding authorisation. The 1992 Further and Higher Education Act instructed the Further Education Funding Council to support courses leading to NVQs and GNVQs. Having said that, Schedule 2A of this act permitted the Secretary of State to authorise funding for other vocational qualifications too. During the mid-1990s, this included BTEC awards, City & Guilds awards, RSA awards, Pitman awards, and others too. An extended quotation from Raggatt & Williams provides useful insights:

The awarding bodies found that their non-N/SVQ and GNVQ products attracted a considerable and continuing demand, and accordingly continued to offer them. [...] BTEC had agreed to phase out its First and National Diplomas (and the part-time Certificates) in favour of GNVQs. Yet demand for these products remained high; indeed it was rising for the National awards. Perhaps unsurprisingly, then, BTEC, which had become independent from the government in October 1993 and was now operating as a more explicitly commercial organization, chose to keep its National Diplomas and Certificates as well as GNVQs. It put up various justifications for this reversal of policy, including the need for more time to effect the changes needed and the claim that GNVQs were not adequate replacements for some of its products [...]

While the government could have made greater use of its powers to remove qualifications from the FEFC's approved list, it was recognized that, because the market was expressing a clear demand for other awards, they clearly possessed attributes that N/SVQs and GNVQs did not.

(Raggatt & Williams, 1999, page 152)

Sharp (1999) noted that, by the mid-1990s, GNVQs and BTECs were attracting students with somewhat different aspirations, quoting statistics that showed that Advanced GNVQ students were considerably more likely than National Diploma students to stay in education, while Diploma students were considerably more likely to progress directly into employment. This provided an impetus to retain the BTEC route. Into the 2000s, the introduction of GNVQ model 4, the AVCE, provided further impetus to continue developing the BTEC National approach, particularly as the style of the new AVCE was now very much closer to the A level approach.

Whereas, at the outset of the 1990s, the NCVQ anticipated that all technical and vocational qualifications would eventually be subsumed within the NVQ framework,

by the mid-1990s it was clear that distinctive awards from the major players – including the BTEC, City & Guilds, the RSA, and others – were continuing to attract large numbers of candidates. During the early-2000s, the 2002 BTEC National suite proved to be very popular, attracting many students back from AVCEs, particularly in curriculum areas like Creative Arts & Media, and Sports (Hodgson & Spours, 2003), as well as in Business (Torrance, Colley, Garratt, et al, 2005). The flight from AVCE also created a new client base for BTEC awards, comprising schools that had never offered vocational courses prior to the introduction of GNVQs.

Evolving qualification models

BTEC awards were conceived very differently from NVQs, being underpinned by a distinctive philosophy of teaching and learning. The following sections explain how TEC and BEC precursors to the CASLO approach were to morph into CASLO qualifications as the BTEC model evolved.

It is important to remember that the BTEC was established as a validating body, to provide national credibility for the work of local colleges. It achieved this through 3 main activities: validation, moderation, and monitoring (see BTEC, 1989). Validation established a ‘contract’ between centre and the Council, to provide reassurance concerning delivery quality. It considered course aims and structure, how the course would be managed, run, and evaluated, and the involvement of local employers and organisations. Moderation was designed to observe this contract in action, providing a quality control function as well as a quality assurance function with effective feedback at its heart. It considered course management, teaching and learning strategies, assessment, course review and evaluation. Finally, monitoring was conceived as a quality audit of processes and decisions. This included analysis and evaluation of the work of centres, and of student achievement, both via random and non-random sampling.

In short, while the Council was responsible for the BTEC qualification model – its structure, content, standards, and approach – centres were responsible for developing BTEC programmes, including both teaching and assessment. In theory, BTECs offered considerable potential for creative course design. In practice, however, even during the early days, it is unclear the extent to which this flexibility was capitalised upon. Fisher went so far as to claim that the opportunity for creativity and freedom may actually have been less than for a typical A level syllabus of its time:

Centres did not at the National level, at least in any meaningful way, have freedom regarding the content and implementation of the curriculum but merely explained to BEC/BTEC how they intended to operationalise that that was very clearly laid down. While centre-devised modules could be written, these were rare

and had to be produced in accordance with specified formats (BEC, 1977c) and were often substantially amended by BEC/BTEC before receiving approval.

(Fisher, 2004, page 248)

Roy Fisher was a history graduate with considerable experience as a lecturer and curriculum developer in post-compulsory education. He was awarded a PhD in 1999 for his investigation into how the BEC-BTEC model evolved over time, studied through the lens of the National award in Business (see Fisher, 2003). His analysis focused upon 3 ‘generations’ of the model, from 1979 to 1992, just prior to the arrival of the Advanced GNVQ in Business. Table 2 is adapted from the appendix to this article and from a similar table in his PhD thesis (Fisher, 1999, page 196).

BEC National Diploma in Business Studies	BTEC National Diploma in Business and Finance	BTEC National Diploma in Business and Finance
Introduced Sept 1979	Introduced Sept 1986	Introduced Sept 1992
Module-based	Unit-based	Module-based
General Objectives + Learning Objectives	General Objectives + Indicative Content	Outcomes + Performance Criteria + Range Statements + Evidence Indicators
6 core modules + 6 option modules	5 core units + 7 option units + business-related skills	8 core modules + 8 option modules + common skills
In-course assignments + internally set end-of-module exams (BEC moderated)	In-course assignments + internally set end-of-unit ‘final assignments’ (BTEC moderated)	In-course portfolio building recording evidence of achievement, including assignments, simulations, work placement tasks, etc. (BTEC verified)

Table 2. Evolution of the BTEC Business National (from Fisher, 2003)

Fisher commented positively on the transition from Generation 1 (BEC) to Generation 2 (BTEC), which reduced the prescriptiveness of the specified learning outcomes (see also Ellison, 1987). He provided an example of an information technology outcome from a Generation 1 module, which listed specific outcomes below a general one:

C Understand the importance of the computer as an information tool and be aware of its impact on administrative operations

C1. describe the main characteristics of the computer, including both hardware and software, recognizing the special need for relevant and accurate input data;

C2. identify the main commercial applications of computers from routine data processing to the provision of management information;

C3. outline the way in which specific administrative procedures have changed in response to the introduction of computer systems.

He contrasted this with essentially the same outcome from the Generation 2 version of the reconfigured module:

D Assess the uses of electronic technology as a means of communication

This single, general outcome from the 1986 module was supplemented by 5 areas of 'indicative content' such as 'main commercial applications of computers for routine data processing' (Fisher, 2003).

Although all 3 generations were structured around learning outcomes of one sort or another, it was the 1992 model that unambiguously embraced the CASLO approach, influenced, of course, by the new NVQ regulations and negotiations with the NCVQ. Fisher's comparable Generation 3 example now read like this:

Outcome 7.3 Assess the applicability of, and where appropriate use, relevant technology in the operation of administrative procedures and systems

Performance criteria

a major applications of technology in administrative operations identified and classified

b factors affecting efficient and effective use of technology assessed

c technology required for particular administrative operations recognized and, where appropriate, used effectively

d introduction of new technology examined and evaluated

Fisher argued that Generation 2 represented the 'Golden Age' of BTECs, with broader objectives and merely 'indicative' content permitting an element of flexibility for staff and students. His analysis of the transition to Generation 3 was damning:

Between 1979 and 1992 further education colleges experienced a period of curriculum implementation and development when learning and teaching styles were transformed from the 'chalk and talk' model to student-centred approaches that were integrated and coherent (for a discussion of this see Fisher, 2003). The influence of the NCVQ in enforcing the instrumentalism of competence would replace this with a fragmented portfolio culture that buckled under the weight of its own monitoring and recording fetish

(Fisher, 2004, page 252)

Carter (2012) described a similar cross-generational progression within Engineering qualifications. This included an extract from a 1986 BTEC National Science Level II unit which echoed the transition to general objectives and indicative content in the Business National. The unit was intended to occupy 60 hours in a part-time course (complemented by work-based learning) or up to 90 hours in a full-time course. The unit was intended to be delivered primarily through practical laboratory experimentation and modelling. It was suggested that the unit might be assessed by practical assignments (50%) and other types of test (50%).

The first element of the unit (which comprised 5%) was specified as follows (the numbers indicated general objectives, and the letters indicated indicative content):

1. Organise elements and information relating to engineering problems by identifying internal and external systems
 - a. system boundary
 - b. sub-system
 - c. interactional paths
 - d. effect of component interaction, eg interpretation of symptoms in fault diagnosis
2. Identify the significant features of systems and represent them by block diagrams
 - a. inputs and outputs
 - b. directions of signal flow
 - c. concept of signal modification and conversion

Systems referred to might include, for example, diesel engine-generator set, air compressor, machine tool, robotic arm, manufacturing coil, or vehicle drive system.

(Carter, 2012, page 226)

BTEC Firsts and Higher Nationals also came to adopt the CASLO approach. This can be illustrated using information from the BTEC Higher National Mechanical Engineering specification that was published for first teaching in September 2000, to be regulated under the new National Qualifications Framework.

These qualifications were offered as either Higher National Certificates or Higher National Diplomas, and both of these courses were designed to be taken over 2 years. Built from units of approximately 60 Guided Learning Hours each, the part-

time HNC required 6 mandatory units plus 4 optional ones, while the full-time HND required the same 6 mandatory units plus 10 optional ones. HNCs were designed for those in work, whom (it was assumed) would be gaining extra experience that would roughly equate to the extra learning time in HNDs (Judith Norrington, personal communication). All units were structured according to the CASLO approach, in terms of both learning outcomes and assessment criteria, as illustrated in Table 3.

Learning Outcomes	Content	Assessment Criteria
<p>1. Select and apply costing systems and techniques</p>	<p>Costing systems: job costing, process costing, contract costing</p> <p>Costing techniques: absorption, marginal, activity-based</p> <p>Engineering business functions: design, manufacturing, engineering services</p> <p>Measures and evaluation: break-even point, safety margin, profitability forecast, contribution analysis, 'what if' analysis, limiting factors, scarce resources</p>	<ul style="list-style-type: none"> • Identify and describe appropriate costing systems and techniques for specific engineering business functions • Measure and evaluate the impact of changing activity levels on engineering business performance
<p>2. Analyse the key functions of financial planning and control</p>	<p>Financial planning process: short, medium, and long-term plans, strategic plans, operational plans, financial objectives, organisational strategy</p> <p>Factors influencing decisions: cash and working capital management, credit control, pricing, cost reduction, expansion and contraction, company valuation, capital investment</p> <p>Budgetary planning: fixed, flexible and zero-based systems, cost, allocation, revenue, capital, control, incremental budgeting</p> <p>Deviations: variance calculations for sales and costs, cash flow, causes of variance, budgetary slack, unrealistic target setting</p>	<ul style="list-style-type: none"> • Explain the financial planning process • Describe the factors influencing the decision-making process during financial planning • Examine the budgetary planning process and its application to financial planning decisions • Apply standard costing techniques and analyse deviation from planned outcomes

Learning Outcomes	Content	Assessment Criteria
3. Apply basic project planning and scheduling methods to a specified project	Project resources and requirements: human and physical resource planning techniques, time and resource scheduling techniques, Gantt charts, critical-path analysis, computer software packages, work breakdown structure, precedence diagrams	<ul style="list-style-type: none"> • Establish the project resources and requirements • Produce a plan with appropriate time-scales for completing the project identify human resource needs • Identify approximate costs associated with each stage of the project

Table 3. Outcomes from the 2000 Mechanical Engineering Higher National Unit 1 (Business Management Techniques)

For each core and optional unit, a content list was provided for each learning outcome. Although units were designed to be free-standing, centres were encouraged to be innovative in designing programmes that enabled integration and flexibility within and across outcomes from different units. All of the units were designed to recognise the importance of skills development through the integration of Common Skills.

Again, in accordance with the CASLO approach, students were required to achieve all of the specified learning outcomes (and assessment criteria) to be awarded a unit pass. Each unit also awarded higher grades, albeit using generic grading criteria rather than unit-specific ones. These were applied to the totality of assessment evidence provided for the unit. Centres were encouraged to incorporate a variety of traditional and innovative assessment methods, including case studies, assignments, time-constrained assessments, and work-based projects.

Whereas qualifications in the Higher National suite tended to grade units on the basis of generic criteria, qualifications from the National suite tended to take a different approach, specifying unit-specific criteria for pass, merit, and distinction. This is illustrated in Table 4 with a grading grid from Unit 1 from the BTEC Edexcel Level 3 in Business, which had been introduced for first teaching from September 2007. To be awarded a merit grade on this unit, students would need to have satisfied all 3 merit criteria and all 5 pass criteria. If they also satisfied the single distinction criterion, then they would be awarded a distinction grade.

Just like the 2000 Higher National in Mechanical Engineering, this 2007 National in Business also included a content list, bespoke to each learning outcome of each unit. Its specification also provided detailed unit-specific guidance for tutors on:

- delivery approaches
- assessment approaches (with criterion-specific tips) ¹⁰⁸
- links to occupational standards, other BTEC units, and other qualifications
- essential delivery resources (access to computers, books, and so on)
- indicative reading for students (textbooks, journals, websites)
- links to key skills

¹⁰⁸ For example: "In P4, explanations can be diagrammatic with suitable annotations as well as written or orally presented. The functional activity should be selected so that learners are able to demonstrate an understanding of the complexity and interdependency of functional areas and how they function in contrasting organisations."

Learning outcomes	Pass criteria	Merit criteria	Distinction criteria
<p>1. Understand the different types of business activity and ownership</p> <p>2. Understand how the type of business influences the setting of strategic aims and objectives</p> <p>3. Understand functional activities and organisational structure</p> <p>4. Know how external factors in the business environment impact on organisations</p>	<p>P1 describe the type of business, purpose and ownership of two contrasting organisations</p> <p>P2 describe the different stakeholders who influence the purpose of two contrasting organisations</p> <p>P3 outline the rationale of the strategic aims and objectives of two contrasting organisations</p> <p>P4 describe the functional activities, and their interdependencies in two contrasting organisations</p> <p>P5 describe how three external factors are impacting upon the business activities of the selected organisations and their stakeholders</p>	<p>M1 explain the points of view from different stakeholders seeking to influence the strategic aims and objectives of two contrasting organisations</p> <p>M2 compare the factors which influence the development of the internal structures and functional activities of two contrasting organisations</p> <p>M3 analyse how external factors have impacted on the two contrasting organisations.</p>	<p>D1 evaluate how external factors, over a specified future period, may impact on the business activities, strategy, internal structures, functional activities and stakeholders of a specified organisation.</p>

Table 4. Grading grid from the 2007 Business National Unit 1 (Exploring Business Activity)

Finally, it is worth noting how the idea of locally-devised programmes – which was the principle that led to the TEC and the BEC being established as validating bodies – gradually fell out of favour over time. For instance, the specification for this 2007 BTEC Business National explained that centres would normally be able to meet local needs by selecting the most appropriate of the specialist optional units on offer. In certain circumstances, they might be able to make a case for incorporating units from other BTEC National specifications. But only in exceptional circumstances would they be permitted to develop their own units. Permission would only be granted on the basis of strong evidence that local needs could not be met using standard units.

Evaluations

The QCA (2005) investigated standards in the 2002 suite of Nationals, focusing on awards in Media, Business, and Maintenance and Operations Engineering. It noted evidence of good and poor practices, but concluded that national standards were being maintained overall with the transition to the new qualifications. The QCA report commented on 2 aspects of qualification design that set BTECs apart from NVQs in their adoption of the CASLO approach. The first was the inclusion of an Integrated Vocational Assignment, an holistic unit that aimed to synthesise learning from multiple units, to help students appreciate “the seamless relationship between units in an applied vocational context” (QCA, 2005, page 17). As for all of the other units, this was assessed by the centre against specified outcomes and criteria, and quality assured via internal and external verification. The second was unit grading, which required criteria to be specified at multiple levels for each unit outcome – pass, merit, and distinction – in the form of a grading grid. In theory, higher criteria were intended to reflect a qualitative improvement in performance, although the report noted that they sometimes required additional tasks to be undertaken, which was not intended. The report recommended a number of steps that (BTEC owner) Edexcel could take to improve the suite, including:

- clearer guidance on grade differentiation, together with a review of units to ensure qualitative rather than quantitative reward of performance
- regional events for internal and external verifiers, to standardise and maintain national standards and to provide contextualised guidance (focused on sufficiency of evidence, assessment design, grading and differentiation)

- good quality exemplar assignment material and specific guidance on how centres can develop the assignment writing skills of their teaching staff ¹⁰⁹

Challenge in interpreting, and differentiating between, grading criteria was a consistent theme in the QCA report, across all 3 qualifications. The percentages of centres experiencing difficulties of this sort were 68% for engineering, 31% for business, and 13% for media. Problems included business centres interpreting the terms 'analyse' and 'evaluate' in different ways, engineering centres expressing concern over vague or unclear criteria, and media centres noting repetition and overlap of criteria across units.

Challenge in interpreting, and differentiating between, grading criteria was also a theme in Ofqual's follow-up monitoring report on the Edexcel Level 3 BTEC National Certificate in Manufacturing Engineering, which had been introduced in September 2007 prior to the introduction of the QCF (Ofqual, 2010a). Improvements were built into revised unit specifications that were being prepared for accreditation to the QCF. They provided a clearer structure to the units, and demonstrated more clearly what was required for learners to achieve a pass, merit, or distinction.

Conclusion

By the turn of the millennium, the CASLO approach had become the high-level design template for BTEC qualifications across all 3 of its principal suites: First, National, and Higher National. Unlike the NVQ approach, however, the 'BTEC way' paid more than lip service to curriculum issues, and BTEC programmes were associated with a strong philosophy of teaching and learning that elevated the role of projects, problem solving, team working, and student ownership of the learning journey.

Although bearing more than passing similarity to the GNVQ model, the BTEC model proved to be far more successful. Ironically, while successive iterations of the GNVQ model became less committed to the CASLO approach, successive iterations of the BTEC model became more committed. When the non-CASLO AVCE was finally withdrawn, much of its market share went to the full-CASLO BTEC National, which then went from strength to strength.

This is not to say that the BTEC model was immune to CASLO-related problems of the sort that beset NVQs and GNVQs. For instance, case studies of particular BTEC Nationals have raised concerns related to the risk of poor-quality teaching and learning associated with the detailed specification of outcomes and criteria (see, for example, Ecclestone, 2010; Hobley, 2016; Carter & Bathmaker, 2017). Likewise,

¹⁰⁹ Activities and resources of this sort might have been more prevalent within earlier iterations of the BTEC model (Judith Norrington, personal communication).

QCA evaluations have raised concerns related to the risk of BTEC standards not being applied consistently. However, by and large, BTECs have not received anywhere near the level of public critique as NVQs and GNVQs received in relation to their adoption of the CASLO approach. On the one hand, this may be at least partly due to a perception that the TEC, the BEC, and the BTEC councils were keen to collaborate with the educational establishment, while the opposite perception seems to have been true of the NCVQ, particularly in relation to NVQs. On the other hand, it is also probably at least partly due to their rollout never having been as shambolic as was the case for both NVQs and GNVQs.

Genesis

The CASLO approach took root in England as the NVQ framework was rolled out. Initially, it was anticipated that all technical and vocational qualifications would be accredited to the new framework and would come to embrace this new approach to qualification design. Indeed, the principal architect of the NVQ system, Gilbert Jessup, anticipated a time when all qualifications would adopt this new approach – technical, vocational, and general alike – which would embed it at the heart of education and training in England.

Although the CASLO approach was ultimately deemed unsuitable for general qualifications, the NCVQ did locate the approach at the heart of a new, middle route qualification, the General NVQ, or GNVQ. Rollout was highly problematic for both GNVQs and NVQs, yet with different consequences. The NVQ model evolved over time, but it remained firmly grounded in the CASLO approach. The GNVQ model also changed over time, but more radically. Having become increasingly general, it was ultimately replaced by the AVCE, which was then replaced by the Applied A level, and all vestiges of the CASLO approach were jettisoned.

The integration of the CASLO approach within BTECs provides for a more subtle and intriguing story. Both the TEC and the BEC had pioneered an outcome-based approach to qualification design during the 1970s. When the BTEC was established, in 1983, still prior to the introduction of NVQs, the new Council continued to promote an outcome-based approach, recognising this as a solution to problems identified with qualifications of the past, including some arbitrariness of syllabus content.

Despite pioneering outcome-based approaches, neither the TEC nor the BEC appeared to require stringent application of the mastery principle. The general idea of mastery was felt to be important, but tricky to operationalise. They also grappled with the challenge of how to pitch learning outcomes at an appropriate level of generality. The earliest TEC specifications presented fairly specific outcomes, while the earliest BEC outcomes were pitched at a slightly higher level of generality.

The most general outcomes appeared within the 1986 BTEC specifications, which applied a common approach across both technician and business awards. Outcomes were now specified at a high level, with each outcome linked to indicative content. The 1986 model also emphasised the importance of criterion-referencing, which seemed to recommend applying the mastery principle more stringently than in previous years. As such, the essence of the 1986 BTEC model was quite similar to the CASLO approach. However, as with the earlier TEC and BEC models, the 1986 model still appeared to permit a certain amount of flexibility in the approach that a college might adopt to criterion-referencing (see BTEC, 1986). For reasons of this sort, we decided not to describe BTECs as the first CASLO qualifications of national prominence.

We do, however, see the CASLO approach unambiguously embedded within the 1992 BTEC model, with its centrally specified learning outcomes and performance criteria, and its stringent application of the mastery principle. Influenced, of course, by NVQ framework accreditation criteria, these BTEC specifications would seem to be the first to bear all of the hallmarks of the CASLO approach. The CASLO approach would soon become one of the defining features of the 'BTEC way' across all 3 principal suites.

Chapter 4. Dominance

The first decade or so of the new millennium was associated with step changes in the regulation of VTQs in England. By the middle of the 2010s, it had become clear that the CASLO approach now dominated the TVET qualification landscape in England. These 2 observations are related, as regulations covering VTQs came to prescribe more and more of the core characteristics associated with the approach.

The NQF

In April 1995, as concerns over NVQs and GNVQs were coming to a head, the Chairman of the School Curriculum and Assessment Authority, Sir Ron Dearing, was invited to consider and advise on ways to strengthen, consolidate and improve the framework of 16-19 qualifications in England, Wales, and Northern Ireland. At the heart of his subsequent report was a proposal for “a coherent national framework covering all the main qualifications and the achievements of young people at every level of ability” (Dearing, 1996, page 3). The intention underlying this new framework was to incorporate existing qualifications – rather than to engineer new ones – and, in particular, to bring the structure of A levels and GNVQs into closer alignment. This was to help secure parity of esteem across these qualifications, as well as to facilitate programmes of learning that incorporated both qualifications. Dearing recommended 3 distinct pathways, differentiated on the basis of their purpose:

1. A level and GCSE – where the primary purpose was to develop knowledge, understanding and skills associated with a subject or discipline
2. applied education (GNVQ) – where the primary purpose was to develop and apply knowledge, understanding and skills relevant to broad areas of employment
3. vocational training (NVQ) – where the primary purpose was to develop and recognise mastery of a trade or profession at the relevant level

These pathways would operate at 3 levels – an Advanced level (corresponding to A level), an Intermediate level (corresponding to GCSE grades A* to C), and a Foundation level (corresponding to GCSE grades D to G). They would be supported by a common Entry Level.¹¹⁰

The Qualifications and Curriculum Authority (QCA) was established by the 1997 Education Act, which was the last Conservative Act to be passed before Labour took

¹¹⁰ Although Dearing promulgated the idea of a formal national qualifications framework, it is worth noting that the NCVQ had developed (and was informally using) essentially the same framework structure some years prior to his report (see Hyland, 1994).

the reins. It gave the QCA power to accredit qualifications, which included developing and publishing accreditation criteria.

The National Qualifications Framework (NQF) was introduced in 2000 as part of the Curriculum 2000 reforms. NQF qualifications were accredited under 1 of 3 broad categories according to their primary purpose: General Qualification, Vocationally-Related Qualification, or Occupational Qualification (see Figure 11, reproduced from QCA, 2000, page 5).

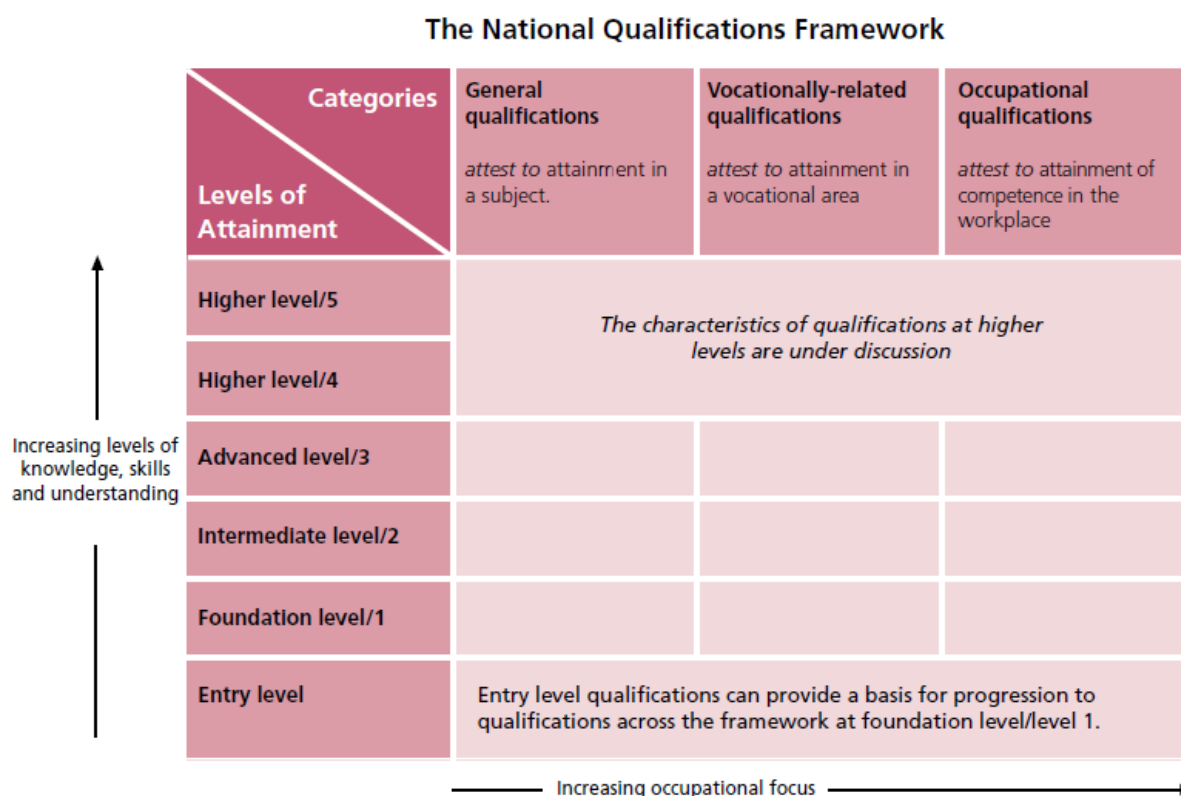


Figure 11. The National Qualifications Framework

Although it was intended that the NQF should incorporate many existing qualifications – suggesting that it was more of a descriptive framework than a prescriptive one – it was deemed essential that all qualifications should satisfy accreditation criteria, to promote transparency, quality, and rigour. These were published in a document entitled ‘Arrangements for the Statutory Regulation of External Qualifications in England, Wales and Northern Ireland’ (QCA, 2000), which was known informally as the ‘Statutory Regulations’.

The Statutory Regulations explained that accreditation criteria consisted of: criteria setting out the required characteristics of qualifications, and criteria – normally in the form of a code of practice – setting out necessary procedural standards. Importantly,

the newly regulated market was not to be restricted purely to GNVQs and NVQs. Any qualification that was submitted for accreditation, and that met the relevant accreditation criteria, would be accredited.¹¹¹ Of direct relevance to the CASLO approach, the Statutory Regulations specified that all Vocationally-Related Qualifications must:

be constructed of units with content expressed as assessable outcomes of learning which provide worthwhile learning goals in their own right

(QCA, 2000, page 17)

In addition, all Occupational Qualifications must:

be directly based on relevant national occupational standards [...]

be constructed of units with content expressed as assessable outcomes of learning

(QCA, 2000, page 20)

The Statutory Regulations also made reference to assessment criteria, although there was no explicit requirement that they should be nested within learning outcomes. Nor was there any requirement that all learning outcomes must be achieved for a qualification to be awarded. However, the very fact that all Occupational Qualifications (OQs) and Vocationally-Related Qualifications (VRQs) had to be specified in terms of units and learning outcomes would seem to acknowledge the growing influence of the outcome-based approach.

When the second edition of the Statutory Regulations was published (QCA, 2004a), there were no longer separate criteria for OQs or VRQs. However, the criteria that were common to all qualifications specified that:

47 A qualification must normally be made up of units that can include a core of mandatory units and a range of optional units, except where the qualification is of an established type that has not historically been unitised, such as the GCSE.

(QCA, 2004a, page 18)

50 The subject matter of the units and/or the qualification as a whole must: [...]

c) be expressed in terms of what a successful candidate will have learned or will be able to do

(QCA, 2004a, page 19)

¹¹¹ Criteria specific to the following were developed: Entry level Qualifications, General Qualifications, Vocationally-Related Qualifications (with additional criteria specific to GNVQs), Occupational Qualifications (with additional criteria specific to NVQs), National Occupational Standards, Key Skills.

So, although expressed slightly differently here, the requirement to specify OQs and VRQs in terms of units and learning outcomes continued. It is worth noting that the Statutory Regulations were quite explicit over quality assurance arrangements, for example:

59 The awarding body must take steps to ensure that internal assessment is carried out in the same way across centres by providing a full assessment specification, including, where appropriate, assessment criteria, mark schemes, exemplar material, and guidance on the use of witness statements.

60 The awarding body must have arrangements in place to enable internal assessors to meet their responsibilities. These arrangements must include, where appropriate, providing assessors with information on:

- a) how to ensure that any tasks set are consistent with the specification;
- b) the nature and type of acceptable evidence;
- c) the extent to which candidates can be allowed to redraft work before it is assessed;
- d) the limits on the assistance that can be given to candidates with work that is to be assessed;
- e) how to ensure that assessment requirements can be interpreted consistently;

(QCA, 2004a, page 21)¹¹²

Credit

Although the NQF Statutory Regulations appear to have played an important role in embedding the idea of learning outcomes within regulated qualifications, it was QCF regulations that fully embedded the CASLO approach, from 2008 onwards. To understand why and how the QCF was introduced, we need to consider the growth of the Credit Movement in England, during the 1990s, which influenced practices in both higher education and further education, albeit in slightly different ways (Pollard, Hadjivassiliou, Swift & Green, 2017). In the further education sector, ideas that stemmed from the Open College Networks, which were subsequently developed by the Further Education Unit, were particularly influential.

¹¹² The previous edition had specified a similar requirement within a section entitled 'Internal assessment' in its common code of practice: "An awarding body must set down assessment criteria, including mark schemes where relevant, to ensure valid and consistent assessment. The awarding body must provide centres with exemplar work showing clearly how defined standards are to be met." (QCA, 2000, page 33).

Open College Networks

As Principal Advisor to the QCA on what was to become the QCF, Peter Wilson was highly influential in its design. He was also an influential figure in the Open College Network (OCN) community – as Co-ordinator of the Leicestershire Open College Network and Chair of the National Open College Network – so it should not be surprising that the QCF was heavily influenced by the OCN approach.

Wilson's detailed account of the emergence of the QCF traces its origins to the development of Open Colleges, Open College Federations, and Open College Networks in different parts of the UK, during the 1970s, which provided alternative progression routes for adults into higher education. The first Open College Network (OCN) credits were awarded in 1983 (Wilson, 2010).

During the early days, these organisations operated locally, without national co-ordination, and credit simply reflected the successful completion of a recognised course of a certain duration at a particular level. With the establishment of the National Open College Network (NOCN) in the late 1980s – and its commitment to develop credit accumulation and transfer agreements during the early 1990s – there was an increased need to reach consensus over what these credits were actually being awarded for (Wilson, 2010). This facilitated the move to an outcome-based approach. Paralleling developments in the NVQ system, the award of credit became more closely associated with the achievement of learning outcomes than with the completion of a programme of study.¹¹³ Yet, according to Wilson, the rationale for this transition was quite different, being grounded in the rights of individual learners, not in the expectations of employers:

programmes [...] should be designed in such a way as to make explicit, and available for public scrutiny, the “hidden” and “intuitive” assumptions being made by teachers about what they expect students to learn/achieve

(Wilson, 2010, page 36)

The most important point to note about the OCN approach was that it was designed to cater for the adult learning sector, with particular reference to the needs of ‘returning’ adult learners:

Not only does the QCF originate in a policy context with a focus on the needs of adult learners, but the ‘credit’ strand of the QCF draws on a long and rich history of recognising the achievements of adult learners in community-based learning,

¹¹³ Ironically, without the ability to cash-in credit for a qualification, the credit accumulation and transfer system was actually pointless for students (Wilson, 2010). Its significance was more symbolic, establishing a national currency for achievement.

in informal adult learning and in contexts totally outside the 'mainstream' development of qualifications during the same period.

(Wilson, 2010, page 2)

Further Education Unit

The credit movement expanded beyond the Open College Network when Conservative Prime Minister John Major initiated work on the modular curriculum in 1991. The Further Education Unit (FEU) was asked to produce advice relevant to the further education sector, which evolved into a blueprint for a credit framework. This was presented in 'A Basis for Credit?' (FEU, 1992).

The report formally integrated 3 key concepts: credit, level, and unit. Furthermore, borrowing the concept of learning outcome from the OCN tradition, the report added the concept of assessment criteria (to parallel the NVQ distinction between elements of competence and performance criteria). The FEU model provided the foundation for a national credit framework for OCNs in 1994, based on credits, levels, and units, although not qualifications (Wilson, 2010). This was entirely consistent with the OCN mission, which was to formally recognise and reward the small steps of achievement that would not be recognised through a system that revolved around end-point certification. Thus, OCNs awarded credits, not qualifications.

The FEU continued developing the idea of credit accumulation and transfer, in an attempt to make it more generally applicable and palatable. The OCNs awarded credits within systems that were specifically designed to be flexible, responsive to local needs, and easily accessible to local organisations, supporting tailored programmes and customised assessment arrangements (Wilson, 2010). Quite explicitly, the idea of being constrained by an overarching qualification specification was anathema to this philosophy. Proposals in the FEU report 'A Framework for Credit' (FEU, 1995a) attempted to bridge the divide between the unit-driven approach of the OCNs and the qualification-driven approach of the exam boards. It did so by making certain of the core concepts more amenable to qualification providers, including the idea of deconstructing qualifications into units with size-related credit values (which watered down the idea of actually awarding credits to learners).¹¹⁴ Thus, the FEU hoped to achieve its goal of establishing:

¹¹⁴ For the OCNs, unitisation meant that programmes were constructed on the basis of units, with personalisation in mind. For the FEU, unitisation simply acknowledged that qualifications could be deconstructed into their component parts.

a post-16 CAT framework encompassing all curriculum and qualifications from key stage 4 of the National Curriculum/adult basic education to post-graduate level qualifications in HE/professional qualifications

(FEU, 1993, page 2)

Of relevance to the direction that CASLO qualifications were beginning to take by the late 1990s, it is worth noting how this 1995 report defined assessment criteria:

Learning outcomes: what a learner can be expected to know, understand and do.

Assessment criteria: statements of more specific learning outcomes.

(FEU, 1995a, page 11)

The idea of assessment criteria as mini learning outcomes seems to be far looser than the idea of performance criteria advocated by the NCVQ. The FEU developed this perspective in supplementary guidance, explaining that criteria should achieve greater specificity by using a specific action verb, content, and qualifiers that make reference to complexity, and/or autonomy, and/or range (FEU, 1995b).

This guidance document also compared pros and cons between the CASLO approach to qualification design (epitomised by NVQs) and the classical approach (epitomised by A levels). Drawing on conclusions from an FEU report written by Alison Wolf (1993), it proposed that a compromise could be struck within the proposed new credit framework:

In order to achieve consistency and effective communication of what learners know, understand and can do, FEU believes that some combination of three approaches is needed:

- written specifications – of learning outcomes, assessment criteria, and level descriptors;
- exemplars – indications of what should be taught and learned; programmes of study, test papers and their analyses, samples of students' work, etc.;
- networking – of unit writers, teachers, examiners and moderators.

The more widely and effectively exemplars and networking are used, the less specific the learning outcome statements or units need to be.

(FEU, 1995b, page 10)

In fact, the report went on to propose that no matter how clearly learning outcomes were expressed, their interpretation for assessment purposes will usually involve both exemplar materials and professional networking.

Policy impetus

Despite following in the wake of these credit system developments, there was no attempt to make credit integral to the NQF.¹¹⁵ It is worth noting, however, that the NOCN had decided (in 1998) to become an awarding organisation and to seek to develop its own credit-based qualifications within the NQF (Wilson, 2010). This made sense given the QCA's aim to recognise all achievements within the NQF with potential (lack of) funding implications for unrecognised providers. To accommodate individual learner needs in a manner that could satisfy funding requirements, the NOCN introduced flexible qualification structures that offered a wide range of unit choices, with rules of combination designed to help ensure a degree of coherence for the overarching qualification (Wilson, 2010). In practice, local OCNs still continued to award unit credits. Officially, though, in terms of NQF recognition, the NOCN only awarded qualifications.

Enthusiasm for the idea of credit waned, in England, towards the end of the 1990s. By way of contrast:

- Northern Ireland continued to develop its Northern Ireland Credit Accumulation and Transfer System project (Cook, 2001)
- Scotland launched its Scottish Credit and Qualifications Framework in 2001 (Gallacher, Toman, Caldwell, Raffe, & Edwards, 2005)
- Wales formally adopted its Credit and Qualifications Framework for Wales in 2002, which was then launched in 2003 (Arad Research Ltd, 2014)

Pressure from Wales, in particular, helped to reignite interest in credit in England. Wales had been developing the idea of credit for many years, which included the Wales Credit and Modularisation Project (later known as Credis). When English awarding organisations were brought into the planning process for a credit framework for Wales, during the late-1990s, they suggested that England would need to be on board with the approach to make engagement viable for them (Jill Lanning, personal communication).¹¹⁶

Policy makers in England responded positively. In the wake of its discussion document 'Success for All: Reforming Further Education and Training' (DfES, 2002), the Department for Education and Skills committed to working with the Learning and Skills Council and the QCA to review barriers to qualification uptake and to explore the feasibility of a credit-based approach. The white paper '21st Century Skills:

¹¹⁵ The 2004 revision of the Statutory Regulations introduced the idea of assigning a 'credit value' although this seems to have been in the FEU (watered down) sense of a size appraisal. The term 'credit value' was not formally defined in the 2004 regulations.

¹¹⁶ Ultimately, the QCF was incorporated as a component of the CQFW.

Realising Our Potential' (DfES, 2003a) subsequently confirmed government commitment to a credit framework:

The consultation on the Skills Strategy has shown widespread support for developing a national credit framework for adults. This is seen as a way of offering the greatest flexibility and responsiveness, with units of qualifications being assigned credit using a standard system. Supporters argue that adult learners can more easily build up units of credit over time towards qualifications, transferring that achievement between different providers if they wish, and having more choice in the units of qualifications they combine. Employers can put together units of qualifications drawn from different sources to form the training programme that best suits their needs.

(DfES, 2003a, page 84)

At this stage, however, government committed only to exploring the idea of a credit framework for adults, acknowledging that credit “frameworks for young people raise quite different issues” (DfES, 2003a, page 85). These were to be considered separately in the light of the forthcoming 14-19 review. Subsequently, the Working Group on 14-19 Reform endorsed the idea of credit for young people, proposing that:

achievement within 14-19 programmes should be certified by diplomas available at the first four levels of the National Qualifications Framework, and using a credit system compatible with that being developed by QCA for adult qualifications.

(Tomlinson, 2004, page 6)

The QCF

The QCA had been exploring the potential for a unitised credit framework since the late-1990s, in conjunction with the Further Education Development Agency (Unwin, 1999). With a new commitment from the DfES, it undertook to consult on the matter. By the end of 2004, the QCA had released a consultation document entitled ‘A Framework for Achievement: Recognising qualifications and skills in the 21st century’ (QCA, 2004c). This included radical proposals to replace the NQF with an entirely new regulatory framework – not simply a framework for adult returning learners, but a framework that could “encompass all formally assessed learners’ achievements outside higher education” (QCA, 2004c, page 3).

Reform

According to the QCA, the new framework would address concerns that the NQF was:

1. too complicated and difficult to understand

2. insufficiently responsive to the needs of individuals and employers
3. insufficiently inclusive of post-16 awards and programmes
4. too procedurally bureaucratic
5. insufficiently inclusive of post-16 (short course) training providers

To solve these problems, the new framework would incorporate a wider range of units from a wider range of unit providers, including customised awards that would meet specific market needs:

Our proposed design for the framework will make it possible for many more employees to gain credit for in-house training. Private training providers that offer high-quality short courses will be able to participate. Outcomes relating to employment sectors or occupations will be driven by the needs of employers.

(QCA, 2004c, page 2)

The fact that it would be credit-based would mean that combinations of units could be accumulated and transferred easily between qualifications and awarding organisations. The consultation explained that all achievements would be structured as units – from which qualifications would be built – and each unit would be defined in terms of: a title, learning outcomes, assessment criteria, a level, a credit value, and a unique database code.

Defining credit as “an award made to a learner in recognition of the achievement of designated learning outcomes at a specified level” (QCA, 2004c, page 19) confirmed that this unit-based framework was conceptually aligned to the 1992 FEU report and to the approach adopted by OCNs.¹¹⁷ As subsequently explained by Ofqual in the introduction to its evaluation of the first 2 years of the new framework, its aim was to provide: “a stable currency for learner achievement across the qualifications system through the award of credit” (Ofqual, 2009b, page 2). The full range of anticipated benefits for learners was summarised in the ‘Final Business Case’ for the QCF:

The QCF offers the opportunity for learners to build up achievements over time and at their own pace. It will allow individuals to achieve smaller packages of learning (units), and, where appropriate, accumulate the associated credits to gain qualifications. No learning will be lost in the QCF, nor will it need to be repeated, as all achievements will be recorded on an individual’s [Learner Record]. Learners will have more control over the routes or pathways that they take through learning, as units can be combined in different ways to meet

¹¹⁷ Bear in mind that, not only was Peter Wilson the Principal Advisor to this programme, the QCA also appointed key figures from the OCN community to lead development teams.

individuals' personal, professional or social needs. Learners will be able to transfer their achievements between all AOs and across all learning providers.

(LSC & QCA, 2008, page 12)

Transition

After a couple of years in development, the new framework – now known as the Qualifications and Credit Framework – was tested and trialled over a 2-year period that lasted from April 2006 to May 2008. Following a decision to proceed, regulations governing the QCF were published by Ofqual in August 2008 (Ofqual, 2008a).¹¹⁸ It was anticipated that all vocational qualifications should be accredited to the QCF by the end of 2010, at which point the QCF would replace the NQF.

Although the QCA had described the QCF as a framework for all achievements, this failed to materialise. It would certainly come to incorporate the vast majority of regulated vocational qualifications. However, certain key qualifications – including GCSEs and A levels – remained outside its orbit, continuing to be regulated under the Statutory Regulations of the NQF. Even the new Foundation, Intermediate, and Advanced Diplomas were regulated outside the QCF.

The situation for NVQs was ambiguous. Wilson argued that the objectives that underpinned the QCF – including simplicity, inclusivity, and responsiveness – effectively undermined the strictures of the NVQ model (Wilson, 2010). QCF regulations permitted existing NVQs to be re-written and submitted for accreditation into the QCF without 'NVQ' in their title (Ofqual, 2008b). Yet, the same regulations also allowed for 'NVQ' to be included in a QCF qualification title as long as the qualification satisfied an additional set of operating rules (Ofqual, 2008b). It would then be regulated as a QCF qualification. In fact, NVQs also continued to exist outside the QCF – as a distinct qualification type – until their regulatory arrangements were finally withdrawn in 2015.

The framework

Across numerous guidance documents, the Qualifications and Curriculum Development Agency (QCDA) explained that the QCF was designed to recognise small steps, enabling students to build up their learning at their own pace,

¹¹⁸ Three months earlier, the QCA had been split into Ofqual, the new regulator for England, and the Qualifications and Curriculum Development Agency (QCDA). The QCDA continued to develop support materials for the QCF (until it was wound up in 2010) while Ofqual focused squarely upon regulation. Technically, Ofqual was still part of the QCA until legislation came into force in April 2010.

accumulating credit that could be built up into a full qualification (see QCDA, 2010a, for example).

The name of each QCF qualification was set out in exactly the same format, to explain how difficult it was (its level), how long it took to study (its size), and what it was about (its content description). This consistency was intended to ensure transparency for anyone who needed to use the information provided by a qualification, for example, an employer making a hiring decision.

There were 9 levels in the QCF, from Entry Level through to Level 8. The lower levels (Entry Level to Level 3) mapped directly onto the NQF. The QCF was also linked to the Framework for Higher Education Qualifications (FEHQ), and higher levels of the QCF (Level 4 to Level 8) mapped directly onto the FEHQ. These linkages are illustrated, below, in Figure 12, which is adapted from QCDA (2010b).

Unit size was expressed in terms of credits, which corresponded to notional hours of learning, enabling qualifications to be classified as either:

- Award (1 to 12 credits – 10 to 120 hours of learning)
- Certificate (13 to 36 credits – 130 to 360 hours of learning)
- Diploma (37 credits or more – 370 or more hours of learning)

Examples of QCF qualification titles included:

- Level 1 Certificate in sport and active leisure
- Level 4 Diploma in buying and merchandising for fashion retail
- Level 8 Award in strategic direction and leadership

Units were accredited to the QCF as either 'shared' (available to all awarding organisations), or 'restricted' (available only to a defined group of awarding organisations), or 'private' (available only to the submitting organisation). The idea of shared units underpinned the principle of Credit Accumulation and Transfer.

Stages of education or employment	QCF levels (and NVQ-NQF levels)		FHEQ levels	
Professional or postgraduate education, research or employment.	Level 8	Vocational Qualifications Level 8	Level 8	Doctoral Degrees
	Level 7	NVQ Level 5 Vocational Qualifications Level 7 Fellowships	Level 7	Master's Degrees Integrated Master's Degrees Postgraduate Diplomas Postgraduate Certificate in Education (PGCE) Postgraduate Certificates
Higher education. Advanced skills training.	Level 6	Vocational Qualifications Level 6	Level 6	Bachelor's Degrees with Honours Bachelor's Degrees Professional Graduate Certificate in Education (PGCE) Graduate Diplomas Graduate Certificates
Entry to professional graduate employment.	Level 5	NVQ Level 4 Higher National Diplomas (HND) Vocational Qualifications Level 5	Level 5	Foundation Degrees Diplomas of Higher Education (DipHE) Higher National Diplomas (HND)
Specialised education and training.	Level 4	Higher National Certificates (HNC) Vocational Qualifications Level 4	Level 4	Higher National Certificates (HNC) Certificates of Higher Education (CertHE)
Qualified or skilled worker. Entry to higher education. Completion of secondary education.	Level 3	NVQ Level 3 Vocational Qualifications Level 3 Advanced Diplomas GCE AS and A Level	N/A	

Stages of education or employment	QCF levels (and NVQ-NQF levels)		FHEQ levels
Progression to skilled employment. Continuation of secondary education.	Level 2	NVQ Level 2 Vocational Qualifications Level 2 ESOL Skills for Life Functional Skills Level 2 Intermediate Diplomas GCSEs at grade A*–C	
Secondary education. Initial entry into employment or further education.	Level 1	NVQ Level 1 Vocational Qualifications Level 1 ESOL Skills for Life Functional Skills Level 1 Foundation Diplomas GCSEs at grade D–G	
Qualifications taken at any age in order to continue or return to education or training.	Entry Level	Entry Level Certificates (1–3) ESOL Skills for Life Functional Skills Entry Level	

Figure 12. Alignment of levels across the National Qualifications Framework, the Qualifications and Credit Framework, and the Framework for Higher Education Qualifications (adapted from QCDA, 2010b, pages 28 to 29).

The CASLO approach

Critical to this section on the dominance of the CASLO approach, QCF regulations now specified all 3 core characteristics as design rules, which meant that units (and qualifications) could not be accredited to the QCF unless they followed the CASLO approach. The blanket nature of this requirement is interesting in the context of the transition away from the approach with the final iteration of the GNVQ (the AVCE) and because the approach was subsequently rejected as a design template for Applied A levels and for the subsequent Diploma qualification. In this context, the lack of evidence of any debate over this blanket requirement seems surprising.

Rules concerning the specification of learning outcomes and assessment criteria were very clear in the new QCF regulations:

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.

(Ofqual, 2008a, pages 11 to 12)

Likewise, the idea of compensation, which underpins the classical approach to qualification design, was formally prohibited by the Regulatory Arrangements:

1.32 All awarding organisations recognised within the QCF award credits and qualifications (see Section 5).

1.33 Credits must be awarded to learners for the successful achievement of the learning outcomes of a unit. The number of credits awarded must be the same as the credit value of the unit. It is not possible for some credits to be achieved for partial completion of a unit or learners to be awarded credit when all the learning outcomes are not achieved by virtue of any 'compensation' for stronger performance in other areas of learning.

(Ofqual, 2008a, page 17)

Beyond Ofqual's regulatory requirements, the QCDA produced a host of guidance documents to help qualification providers comply with the new QCF regulations. For instance, one guidance document, on assessment, included answers to questions like "Do all assessment criteria have to be met for credit to be awarded?" (QCDA, 2010c, page 8).¹¹⁹ Another (QCA-developed) guidance document provided advice on how to articulate learning outcomes and assessment criteria within unit specifications (QCA, 2009). This included both tips, such as how many assessment criteria to write for each learning outcome, and warnings, such as the need to avoid compound statements.¹²⁰ By way of illustration, this document recommended re-writing the compound statement "know about computer hardware, software, and associated health and safety issues" as 3 discrete learning outcomes:

- know about computer hardware
- know about computer software
- understand the health and safety issues associated with the use of computers

Another guidance document discussed the issue of when to construct units primarily from knowledge-based outcomes, or from skills-based outcomes, or when to combine them (QCDA, 2010d).

¹¹⁹ Ironically, although the answer to this question confirmed a clear expectation that each learning outcome must be judged on evidence related to all of its associated assessment criteria, it was a little ambiguous over whether this meant that all assessment criteria actually had to be met for credit to be awarded. The guidance in QCDA (2010e, page 10) was clearer: "The learner must be able to demonstrate all of the assessment criteria for the judgement to be made that the learning outcome has been achieved."

¹²⁰ This guidance was continually being updated. For example, QCDA (2010e) was the fourth version of guidance on how to write units.

Respond to customer requests for repairs L34

LO 1 Know the organisation's housing stock and possible defects which require repair.

AC 1.1 Describe the types of properties which the organisation manages.

AC 1.2 Identify, using the appropriate terminology, the types of faults which can occur in these properties.

LO 2 Know organisational policies and procedures relating to requests.

AC 2.1 Describe the different types of service agreements with customers.

AC 2.2 Identify the organisational policies and procedures relating to repair requests.

AC 2.3 Identify repairs which require emergency action.

LO 3 Be able to establish and respond to customer requests for repairs.

AC 3.1 Deal courteously, sensitively and fairly with individuals.

AC 3.2 Clarify requests from customers to determine the exact nature of what is required.

AC 3.3 Accurately record the details of customers and their requests.

AC 3.4 Identify the other parties involved in the maintenance and repair of the organisation's properties and whether there are any associated charges.

AC 3.5 Identify requests which are outside the organisation's responsibilities.

AC 3.6 Explain how to refer customers to other organisations and individuals.

AC 3.7 Arrange inspection visits and repair work according to organisational procedures and policies.

AC 3.8 Prioritise urgent repairs.

Figure 13. Example of how to combine knowledge- and skill-based outcomes

Figure 13 reproduces an example of when this document considered it appropriate to combine learning outcomes, that is, where "there is a clear relationship between knowledge and skills" (QCDA, 2010d, page 9). In this instance, the assumption is that the first 2 knowledge-based outcomes are required to underpin the third skills-based outcome. While there was no obligation for learners to achieve the learning outcomes in any particular order, the document added, a chronology of learning outcomes where knowledge is followed by action can make the unit look more

coherent. Incidentally, this document also noted that combining knowledge-based and skills-based outcomes can be useful in supporting an holistic and integrated assessment approach:

Where knowledge and skills are separated into different units, there is always the potential that one or the other may be lacking. The advantage of combining the two in a unit for learners is that they can satisfy the requirements to be competent to carry out the function in their job role and can demonstrate to an academic institution that they have the necessary underpinning knowledge and understanding. Designing units with this in mind is more likely to encourage a more holistic and integrated approach in the design of assessment activities for the unit.

(QCDA, 2010d, page 23)

Challenges

The guidance document on writing QCF units is interesting, in retrospect, for the way that it anticipates fundamental implementation challenges for the QCF. Annex D of 'Guidelines for Writing Credit-Based Units' (QCDA, 2010e) concerned how to develop units at Levels 4 to 8, exploring reasons why it might be more difficult to comply with QCF requirements when writing units at higher levels. This included problems with specifying standards (and the challenge of levelling) and problems with assessing standards (and the challenge of testing). In retrospect, it is fair to say that challenges like these threatened the credibility of the QCF in general and not simply at higher levels.

Specification

One set of challenges arose from the requirement that standards for all QCF units had to be specified in terms of both learning outcomes and assessment criteria. This proved to be less problematic when units were derived from older-style National Occupational Standards, typically related to lower-level jobs. However, particularly for units at higher levels, which were often based upon different kinds of standards set by professional bodies or associations, this necessitated a complex process of translation into the QCF format. Even when based upon NOS, it was no longer possible simply to 'cut and paste' statements into the QCF format, as many sectors had departed from the older-style approach to writing standards by then, often failing to articulate critical criteria (QCDA, 2010e). Consequently, even when utilising NOS, unit writing was far from a trivial process.

Levelling

During the early 1990s, when functional analysis was the recommended approach to developing NOS, the degree of challenge associated with any particular NVQ would have been no more nor less than the occupational standard itself, that is, the degree of challenge associated with performing the role adequately. NVQs were assigned to a particular level in the framework, but this levelling was more nominal than substantive, linked to historical hierarchical distinctions between roles, such as between 'technician' and 'craftsman'. It was not substantively important, within the NVQ framework, to be able to infer that all NVQ units at a particular level represented the same degree of challenge.

With the introduction of the NQF, and particularly with the introduction of the QCF, the idea of levelness and of levelling became far more fundamental. Indeed, establishing a degree of challenge for each unit – its level – was crucial to the underpinning logic of the QCF as a credit-based framework. This was because the combination of the specified challenge (level) and the specified size (credits) determined the currency of each unit within the system. This, in turn, was used to justify claims concerning the exchangeability of units within a framework premised upon being able to mix and match units to form bespoke qualifications.

Consequently, the QCF (more so than the NQF) required a mechanism by which the currency of each unit could be specified and verified. Level descriptors, which articulated 3 dimensions of competence for each QCF level, were fundamental to this mechanism. They were published as Annex E of the QCF Regulatory Arrangements (Ofqual, 2008a). The descriptors for Level 2 are reproduced below, for the purpose of illustration:

Summary

Achievement at Level 2 reflects the ability to select and use relevant knowledge, ideas, skills and procedures to complete well-defined tasks and address straightforward problems. It includes taking responsibility for completing tasks and procedures and exercising autonomy and judgement subject to overall direction or guidance.

1st dimension – knowledge and understanding

Use understanding of facts, procedures and ideas to complete well-defined tasks and address straightforward problems.

Interpret relevant information and ideas.

Be aware of the types of information that are relevant to the area of study or work.

2nd dimension – application and action

Complete well-defined, generally routine tasks and address straightforward problems.

Select and use relevant skills and procedures.

Identify, gather and use relevant information to inform actions Identify how effective actions have been.

3rd dimension – autonomy and responsibility

Take responsibility for completing tasks and procedures Exercise autonomy and judgement subject to overall direction or guidance.

These benchmarks were designed to enable awarding organisations to identify an appropriate level for each unit. AOs were expected to compare the learning outcomes and assessment criteria that had been written for each unit against adjacent level descriptors, applying a ‘best fit’ principle to find the best match (Ofqual, 2008a). Of course, this assumed that learning outcomes and assessment criteria had already been written in a manner that appropriately captured the intended degree of challenge for each unit, which made the process somewhat circular.

According to the QCDA, the language used in formulating outcomes and criteria was crucial to defining and communicating the level of a unit. This language should be capable of conveying the appropriate level without reference to a targeted group of learners or to an anticipated context of learning (QCDA, 2010e).

If the language used to write criteria could somehow go beyond the unit-specific details of each outcome to support comparison on a more generic basis, then this would certainly be helpful in warranting claims of unit comparability and, by extension, unit exchangeability. Benjamin Bloom’s Taxonomy of Educational Objectives (which we discussed in chapter 2) was seen as a solution to this problem – in particular, his hierarchical taxonomy of objectives for the cognitive domain, which were ordered as follows:

1. knowledge (lowest level of complexity)
2. comprehension
3. application
4. analysis
5. synthesis
6. evaluation (highest level of complexity)

The QCF levels were based on the assumption that qualitatively different learning outcomes, from qualitatively different units, could be equated (roughly) with

reference to the degree of cognitive complexity of their associated assessment criteria. In other words, units that were written using criteria of a similar level of complexity to ‘analysis’ – that is, criteria implying more complexity than ‘application’ but less complexity than ‘synthesis’ – could be considered to be at the same level. Which level would be determined by reference to the overarching QCF level descriptors.

This logic suggested that units with essentially the same learning outcomes could be written at multiple levels, differentiated only in terms of the level of complexity associated with their assessment criteria. Of particular relevance, here, was the command verb chosen to help determine the levelness of each assessment criterion, with more cognitively complex command verbs being selected for higher levels.

The approach is illustrated in Figure 14 (which is reproduced from QCDA, 2010e, page 31). Note how the command verb chosen for Level 1 (list), is less cognitively complex than the command verb chosen for Level 2 (identify), which is less cognitively complex than the command verb chosen for Level 3 (explain).

Unit title: Understanding health and well-being

Learning outcome The learner will: Understand the political and social context of health and well-being			
Level	1	2	3
Assessment criteria	List the government priorities for health promotion and health education	Identify the main points in government policies to improve the effectiveness of the NHS, especially in relation to preventative health care and health education	Explain the government thinking on how to improve the effectiveness of the NHS, especially in relation to preventative health care and health education

Figure 14. A suggested hierarchy of complexity of command verbs.

The QCDA guidance document went on to illustrate the association between command verbs and levels via Figure 15 (which is reproduced from QCDA, 2010e, page 34). This clarified that there would not be a one-to-one mapping between

command verbs and levels.¹²¹ Indeed, the document was clear that levelness was not determined by any particular command verb, since the meaning of a criterion depends on all of the words used to express it, not just the command verb.

Entry level 3	Level 1	Level 2	Level 3
Define	Define	Apply	Analyse
Demonstrate	Demonstrate	Assess	Apply
Give examples	Give examples	Classify	Clarify
Identify	Identify	Compare	Classify
Indicate	Indicate	Define	Critically compare
Locate	Locate	Demonstrate	Demonstrate
Outline	Outline	Describe	Develop plan/idea
State	State	Differentiate	Diagnose
Use	Use	Distinguish	Differentiate
		Estimate	Distinguish
		Give (+/-points)	Draw conclusions
		Illustrate	Estimate
		Perform	Evaluate
		Select	Explain
		Use (a range of...)	Extrapolate
			Implement
			Interpret
			Judge
			Justify
			Perform
			Review and revise
			Summarise

Figure 15. The (loose) association of command verbs with unit levels.

¹²¹ Note, for instance, that 'demonstrate' appears in all 4 columns and 'define' appears in 3 columns.

Having said that, it is also important to recognise that certain command verbs did not appear in this table until higher levels, so their incorporation within assessment criteria might be used to indicate that the unit might well be pitched at a higher level.¹²² It was exactly this (loose) association between command verb and level that provided a rationale for using them to help convey intended degree of challenge.

It is worth noting that all 3 of Bloom's most complex categories of cognition – analysis, synthesis, and evaluation – appear among the command verbs listed for Level 3. This begs the question of how to write effective criteria for Levels 4 to 8, which brings us back to the guidance document on writing QCF units that anticipated fundamental implementation challenges for the QCF:

One possible impact of this shifting focus across different levels is that it becomes more difficult to develop precise and easily measurable learning outcomes and assessment criteria at higher levels of achievement. There is a danger that assessment criteria at higher levels either become repetitive, or that they fail to establish an explicit assessment standard for the unit.

(QCDA, 2010e, page 58)

[...] evidence from the QCF test and trial programme suggests that it is actually more difficult to establish meaningful distinctions between units at higher levels of the framework. [...] It is possible that the reason why it becomes practically more difficult to distinguish between levels of achievement as one proceeds up the levels of the QCF is that the distinctions between levels 4 to 8 are theoretically less easy to establish. Although the levels of the QCF are nearly always presented as a neat and even set of 'stages' in a hierarchy of achievements, perhaps in reality these stages get progressively 'narrower' as one goes up through the levels. The difficulty in identifying the difference between a unit at level 6 and one at level 7 may actually be a reflection of reality.

(QCDA, 2010e, page 59)

Whereas the hierarchical structure of objectives appeared to provide a rough-and-ready solution to the challenge of defining and communicating levelness for lower-level units, the same does not seem to have been true for higher-level ones. There appears to have been genuine concern over the potential to capture the levelness of a unit via appropriately worded learning outcomes and assessment criteria.

Finally, it is worth mentioning a criticism that is sometimes voiced, but not well documented, related to the use of command verbs within QCF units. Problems arose when unit writers, who had been charged with developing qualifications at a

¹²² For example, 'estimate' appears in the Level 2 and Level 3 columns, while 'analyse' and 'evaluate' only appear in the Level 3 column.

particular level, interpreted the link between command verbs and levelness far more strictly than can possibly be justified (as though the use of certain command verbs within criteria for a unit straightforwardly warranted the claim that it was pitched at a certain level).¹²³

Imagine, for instance, that we treated 'analysis' as a skill that resides inherently at Level 3, suggesting that when we see analysis occurring, we can safely infer that the analyser is operating at Level 3. If this were true, then there would be some legitimacy in treating units with assessment criteria framed in terms of command verbs like 'analyse' as though they were comparable and, therefore, exchangeable. If so, then command verbs would provide a simple – and easily verifiable – tool for engineering the kind of comparability required by the QCF. Unfortunately, this is not true, not even roughly. Command verbs are not that definitive. As recognised in the QCDA guidance, command verbs alone are insufficient to define the degree of challenge associated with a unit. Sometimes 'analysis' will be associated with a very low level of challenge, other times with a very high level. There is more to defining and communicating degree of challenge than can be captured by the use of a particular command verb. Exactly what that might be was assumed to be part of the body of expertise required to be a competent QCF unit writer.

Assessment

The QCDA guidance document on writing QCF units also grappled with the challenge of developing assessment methods at higher levels, where learning outcomes were often framed mainly in terms of knowledge and understanding, and especially for qualifications that had traditionally relied exclusively upon tests or exams. The document was clear that this was acceptable, in principle, although:

the requirements for assessment of such units are exactly the same as for other levels of the QCF: all the learning outcomes of the unit must be achieved to the required assessment standard in order for credit(s) to be awarded for that unit.

(QCDA, 2010e, page 58)

This raised (rather than answered) a fundamental question concerning the nature of testing within the QCF, which appears never to have been fully resolved. The question arises because written tests and exams tend to be associated with the classical approach to qualification design, which is based upon a compensatory (rather than a mastery) aggregation principle, and which is therefore far more open to sampling across learning outcomes. Under the QCF, not only did learners need to provide evidence of having achieved each outcome, they actually needed to provide

¹²³ Ofqual's report on grading VTQs provided some evidence related to this criticism (Newton, 2018). Additional insights were provided by Barry Smith and Norman Gealy (personal communication).

evidence of having achieved each criterion for each outcome. First, this means no sampling. Second, this means that the assessment of each individual criterion needs to be sufficiently reliable in its own right. This contrasts with the classical approach to assessment design, where sufficient reliability only needs to be demonstrated at the highest level, that is, at the level of the total mark achieved across all qualification components.

Written testing

The dilemma for written testing within the QCF was that, to do it properly – consistent with QCF regulations and the underpinning mastery model – a mini test would need to be created for each criterion of each outcome. Furthermore, each student would need to pass the relevant mini test, for each criterion of each outcome, to pass the unit. It is certainly possible to imagine an assessment designed like this. However, it would probably end up as a mega test, with very many individual test items, and it would probably fail most if not all of the candidates who sat it (as sustaining that level of performance across a very long test is a very big ask). This, of course, is why the CASLO approach tends to be associated with continuous, or staggered, centre-based assessment, not written tests or exams.

Separate guidance on assessment confirmed that it was, in fact, permissible to assess units through tests and exams, for example:

The examination questions must be designed in such a way as to enable an assessor to make an assessment judgement about whether or not the learner has achieved the outcomes of a unit, but the questions themselves may be much more explicit than the learning outcomes of the unit. Indeed, where learners are offered a choice of examination questions, this separation of learning outcomes and examination questions will be essential.

(QCDA, 2010c, page 6)

Yet, it provided no guidance on exactly how this could be achieved without sampling or compensation. Note that the same document specifically ruled out both sampling and compensation. Finally, its guidance on using pass marks is worth reproducing in full:

9. Can we set a percentage pass mark for a unit? If so, what should it be?

Yes, but the percentage pass mark must relate to all the learning outcomes of a unit. A learner cannot be awarded credit for a unit if only a proportion of the learning outcomes have been achieved. Again, the separation of test or examination questions from the learning outcomes of a unit enables such assessment judgements to be made easily.

In effect, a percentage pass mark reflects the level of confidence that the assessor has in the outcomes of the assessment process, not the proportion of the learning outcomes that a learner has successfully completed. The QCF sets no particular requirements about pass marks in such circumstances. Providing the assessor is confident that all the learning outcomes of a unit have been achieved against the stated assessment criteria, credit can be awarded for the unit.

(QCDA, 2010c, page 7)

It is hard to understand what this passage means. But, it seems not to shed any light on the fundamental question concerning the nature of testing within the QCF. Confusion persisted.

Subsequent investigations by Ofqual occasionally surfaced this tension, for example, it arose within a thematic review of Level 6 and 7 qualifications that were available to international students on a 'Tier 4' study visa. The report observed that some awarding organisations used "a compensatory system of assessment for QCF qualifications, in breach of QCF arrangements" (Ofqual, 2014a, page 21):

This issue seems to stem from awarding organisations moving long-running qualifications onto the QCF, but continuing to use compensatory models of assessment (for example, a written exam with a 40 per cent pass mark).

(Ofqual, 2014a, page 21)

Recommendations from the report confirmed that all QCF qualifications must require all learning outcomes to be met for a pass to be awarded, but failed to explain how this might be accommodated within written tests or exams. Finally, the report also expressed concern that command verbs in assessment criteria were "not sufficient for qualifications at level 6 or 7" (Ofqual, 2014a, page 17), illustrating the other key challenge, of levelling.

Questionable transitions

Consistent with its ambition to rationalise the qualifications landscape, the QCA anticipated that all regulated qualifications would become part of the QCF. Pressure to incorporate GCSEs and A levels was resisted, but other long-standing, well-respected qualifications were forced to transition. This included graded performance exams, which had existed for well over a century, including graded exams in music, dance, speech, and drama.

These exams had always been, and still are, fairly unusual in various respects. For instance, they are unusual in terms of target cohort, in being almost entirely elective – adults and young people choose to take them when ready, and typically fund themselves. In terms of qualification design, they are unusual in being based on a

progressive mastery model – learners progress hierarchically up a suite of (normally) 8 qualifications, demonstrating skills of increasing technical difficulty and complexity. They are used for an unusually wide range of purposes, too, from building confidence and self-esteem, to satisfying a personal hobby, to developing technical (occupational) competence.

This family of qualifications is described in a report by Rachael Meech, which is particularly interesting for its account of how they were forced to adapt to increasingly stringent regulatory requirements, especially the QCF (Meech, with McBirnie, & Jones, 2018). She noted that:

Many awarding organisations found elements of the new QCF regulatory criteria challenging to meet whilst simultaneously preserving the ethos and purpose of their graded examinations and associated processes. Organisations faced challenges with the conceptual framework of the QCF and how it would work in practical terms for their units and qualifications, in particular:

- Requirements for centre approval when a centre-based model of assessment is not operated for graded examinations
- Drafting procedures for units and rules of combination for qualifications which were already well-established, having evolved through a set of detailed syllabuses. This procedure proved challenging due to the requirements placed by Section 1 of the regulatory criteria which arguably undermined well understood processes already instituted.

(Meech, et al, 2018, pages 15 to 16)

These awarding organisations struggled to redevelop their qualifications in ways that complied with QCF regulations, while preserving the ethos and value of graded exams. This included having to square requirements for detailed learning outcomes and assessment criteria, plus mastery aggregation, with a well-established approach to assessment based upon compensation and a one-off external exam. Of course, it was not possible to square this circle, which suggests that it was never reasonable to expect to be able to regulate these qualifications effectively under the QCF.

Evaluations

In July 2008, Ofqual committed to evaluating its new regulatory arrangements, to determine whether they were supporting effective regulation. Evaluations of the first and second year of operation – based on QCF user feedback, scrutiny of sampled units, and outcomes from regulatory activities – were published in January 2010 (Ofqual, 2010b) and May 2011 (Ofqual, 2011b).

The interim report (2009 evaluation) expressed satisfaction that 57% of sampled units were fully compliant, with only 10% requiring immediate review, and 32% having only relatively minor technical issues. These issues included mismatch between learning outcomes and assessment criteria, the clarity with which learning outcomes and assessment criteria were expressed, and manageability concerns related to the number of learning outcomes and assessment criteria within certain units.

Although the final report (2010 evaluation) found “broad support” for the framework (Ofqual, 2011b, page 12), it also recognised that some users felt that the QCF model and its regulatory arrangements were “fundamentally flawed” (Ofqual, 2011b, page 12). Ofqual accepted that the speed and nature of the introduction of the QCF had caused problems, acknowledging that many users who commented had criticised the “attempt to apply one set of design rules to a wide range of very different qualifications” (Ofqual, 2011b, page 12). There was concern that QCF design features were more suited to certain qualifications, students, and sectors than to others.

Even after 2 years of operation, it had become clear that opportunities for credit accumulation and transfer were lacking and, more importantly, that learners themselves had not yet expressed significant demand for it. Ofqual was frank that it needed to consider whether the benefits of regulatory requirements that provided for infrequently used flexibilities outweighed the costs that might be incurred. One particular threat arose from shared units being assessed in different ways, reducing the likelihood that comparability of standards could be maintained. A specific concern was that units devised by organisations with specialist expertise (for example, in conflict management training) could be used by organisations that lacked it and failed to understand how it ought to be assessed, presenting a major threat to standards.

In addition to cost and burden challenges related to having to assess all learning outcomes and assessment criteria, the final report also noted that qualifications that had traditionally been compensatory did not fit well into the QCF, particularly higher-level professional qualifications. Grading challenges were also noted. The report ended with a list of lessons learnt:

- the potential pitfalls of a central policy that drives awarding organisations to redesign a large number of qualifications, with a range of different characteristics and purposes, to conform to one set of design requirements
- the need for clarity in the lines of accountability in qualifications design, approval and delivery
- the need for credit, which is the ‘currency’ of units, to be assigned consistently

- the risk that detailed and/or poorly understood regulatory requirements can distract from, or overshadow, more important regulatory principles
- the implications for commercial and/or competing organisations of sharing units and of collaborating with others to develop units; and
- the challenge of imposing titling rules that do not align with established and understood titles.

(Ofqual, 2011b, pages 23 to 24)

Structure versus quality

Although the QCA anticipated that QCF regulations would ultimately displace NQF ones, this did not actually happen. Various qualification types continued to be regulated under the 2004 NQF Statutory Regulations until these were eventually displaced by Ofqual's General Conditions of Recognition (GCR), which were first published in May 2011. When the GCR came into force, the 2008 QCF Regulatory Arrangements document was designated a subsidiary Regulatory Document (applying only to QCF qualifications), as was the 2006 NVQ Code of Practice (applying only to NVQs).

Bearing in mind the goals of the QCF reform process – to enhance simplicity, inclusivity, and responsiveness – it is worth considering whether anything important was lost when transitioning across successive regulatory documents. A particular question arises over the change in emphasis that occurred with the introduction of QCF regulations, and whether this elevated concern for unit and qualification structure over concern for assessment quality.

To be clear, it is not true that QCF regulations ignored assessment quality or took it for granted. For instance, they specifically stated that an awarding organisation must have in place the necessary systems, procedures and resources to ensure:

- a. assessment instruments and tasks are produced to the required quality standards
- b. assessment evidence produced by learners is authentic
- c. accuracy and consistency of standards in the assessment of units, across units and over time [and so on]

(Ofqual, 2008a, page 27)

However, it is true that QCF regulations focused primarily on structural design requirements. Furthermore, the plethora of QCDA guidance documents also prioritised unit and qualification structure over assessment quality. To some extent, this was a deliberate strategy. The new regulations were intended to focus on

expectations of awarding organisations (regarding systems, procedures, resources, and so on) rather than expectations of practices. Hence, there was no QCF code of practice to specify, for instance, the methods by which learning outcomes ought to be developed or assessed.¹²⁴ Indeed, because the QCF recognised such a wide variety of qualifications, without specifying any distinct sub-types, it may not actually have been possible to produce regulations with that level of specificity.¹²⁵

With this shift in emphasis, many of the detailed requirements that had been built into earlier regulations (to prevent or mitigate serious problems that had occurred) no longer featured explicitly in QCF regulations. This included, for instance, the requirement from the Statutory Regulations that an awarding organisation must take steps to ensure that internal assessment is carried out in the same way across centres, by providing support materials such as assessment criteria, mark schemes, exemplar material, and guidance on the use of witness statements (QCA, 2004a). Also no longer featuring explicitly in QCF regulations were requirements from the NVQ Code of Practice, such as the following, from sections headed 'Internal verification' and 'External verification' respectively:

56. Guidance produced by the awarding body must include exemplars of:

- procedures for standardising assessment so that assessors are operating to the same standard
- models for developing an internal verification sampling plan appropriate to the centre's level of assessment activity. Models must ensure that over time all assessors, all assessment methods and all candidate units are included in the sample
- procedures for standardising the judgements and decisions of internal verifiers operating in a centre
- the types of records a centre must keep to demonstrate the effectiveness of its internal verification procedures.

¹²⁴ Note that there were structural rules for writing and assessing learning outcomes. For instance, learning outcomes and assessment criteria needed to be written in a particular format, and judgements needed to be aggregated in a particular way. However, there were neither rules nor guidance on how to derive a set of outcomes for any particular domain of learning (which contrasts with the requirement to use functional analysis for NVQs during the early 1990s). And QCDA guidance was quite explicit that flexibility built into QCF regulations even permitted centres to use different assessment methods for the same unit offered to different groups of learners (QCDA, 2010c).

¹²⁵ Regardless of whether this may also have been undesirable, from the perspective of enhancing flexibility.

(QCA, 2006, page 15)

60. Awarding bodies must require external verifiers to:

- confirm that centres continue to meet the centre approval criteria
- recommend the imposition of appropriate sanctions on centres that fail to meet the requirements
- confirm that assessments are conducted by appropriately qualified and occupationally expert assessors
- sample assessment decisions to confirm that they are authentic and valid and that national standards are being consistently maintained
- confirm that assessment decisions are regularly sampled, through internal verification, for accuracy against the national standards
- check that claims for certification are authentic, valid and supported by auditable records
- [and so on]

(QCA, 2006, page 15)

By way of contrast, requirements in the 2008 QCF Regulatory Arrangements were far less detailed, for example:

The awarding organisation must ensure that it has arrangements in place for standardisation and quality assurance of assessment outcomes across centres and awards.

(Ofqual, 2008a, page 28)

Bearing these considerations in mind, it is reasonable to conclude that assessment quality took a back seat while the QCF regulations were being introduced, as unit and qualification structure took centre stage.¹²⁶ Of course, many awarding organisations would have continued to implement quality assurance arrangements of

¹²⁶ It is worth noting that the 2001 NVQ Code of Practice was even stronger on the provision of exemplar work for standardising internal assessment: “The awarding body must provide centres with exemplar work showing clearly how defined standards are to be met.” (QCA, 2001, page 27). The use of “where appropriate” alongside similar requirements in the 2004 Statutory Regulations (and no comparable requirement in the 2006 NVQ Code of Practice) suggests that the requirement to provide exemplar work for all internal assessments may have proved to have been too stringent. Note that the rationale for the 2006 Code of Practice revision was to focus on quality assurance rather than quality control and to: “reduce perceived bureaucracy and allow the controlled development of innovative ways of assessing and quality assuring NVQs” (QCA, 2006, page 1).

the sort embodied in earlier regulations. Yet, the fact that they were not explicitly required to do so by QCF regulations left a door open to organisations who saw the potential to cut costs by not implementing important controls. Furthermore, this lack of explicit requirement might also have made it harder to justify control-related burdens to centres.

Dominance

Even though the QCF never subsumed all regulated qualifications, it did come to dominate the qualifications landscape in England (until it was withdrawn).

Qualification Type	2010 to 2011	2011 to 2012	2012 to 2013	2013 to 2014	2014 to 2015
Qualifications & Credit Framework (QCF)	54 %	62 %	71 %	78 %	85 %
Vocationally Related Qualification (VRQ)	13 %	10 %	7 %	5 %	3 %
National Vocational Qualification (NVQ)	9 %	7 %	5 %	3 %	1 %
Occupational Qualification (OQ)	1 %	0 %	0 %	0 %	0 %
Other General Qualification (OGQ)	4 %	4 %	3 %	3 %	3 %
A level	2 %	1 %	1 %	1 %	1 %
GCSE	4 %	3 %	3 %	2 %	2 %
Functional Skills	1 %	1 %	1 %	1 %	1 %
All other qualifications	13 %	11 %	9 %	7 %	5 %
Total number of available qualifications	18,095	20,500	23,642	24,965	24,520

Table 5. Number of available qualifications as a percentage of the regulated qualification market

The figures in Table 5 were computed from data taken from the ‘Annual Qualifications Market Report’ for 2014/15 (Ofqual, 2016). The percentage values were calculated in relation to the total number of qualifications that were available to learners, each year, from 2010 to 2015, combined across England, Wales and Northern Ireland. They illustrate how QCF (CASLO) qualifications became more and more dominant, increasing from 54% of the market (during the 2010 to 2011 academic year) to 85% (during the 2014 to 2015 academic year). Where the percentage of other qualification types decreased over time, this was largely a

consequence of transitioning into the QCF. Note that, even in 2015, there were CASLO qualifications that were regulated beyond the QCF across multiple qualification types (including NVQs, VRQs, OQs, VRQs). In other words, we can be confident that, by 2015, the CASLO approach completely dominated the regulated qualifications landscape.

Having said that, it is important to conclude by acknowledging that the approach might already have become dominant in the landscape some time prior to the introduction of the QCF (although it is hard to tell for sure). Bearing in mind the growing significance of BTECs in the regulated market – which had embraced the CASLO approach since the early-1990s – it is quite possible that the approach may have become dominant much sooner, perhaps under the NQF and maybe even earlier. The QCF regulations give us surety that the approach had come to dominate by the mid-2010s, but many qualifications (not just NVQs) had come to adopt it long before these regulations came into force.

Having now considered nearly 3 decades of CASLO qualifications – from the late-1980s to the mid-2010s – we are in a good position to begin unravelling the multiplicity of goals that drove NVQ designers, and designers of other qualifications, to adopt the CASLO approach. The next chapter is devoted to this challenge.

Chapter 5. CASLO goals

Now is a good time to reflect on motivations for adopting the CASLO approach as a high-level design template for qualifications as distinct as NVQs and BTECs, let alone as the design template for all QCF qualifications, or even all qualifications, *per se* (as both Jessup and the FEU appeared to propose). These motivations – the goals that explain why the CASLO approach was adopted in the first place – were not always clear. This is partly because the purposes for which these qualifications were designed were articulated differently by different commentators. But it is also because the motivations that drove adoption of the CASLO approach would have been a subset of the broader profile of purposes proposed for the qualification in question (and were not always clearly distinguished).

Because they determine the design of the qualification model itself, CASLO-specific goals would have been extremely fundamental. Yet, perhaps for exactly that reason, the need to articulate them explicitly may not always have been apparent. Whatever the explanation might be, it seems that the goals that drove adoption of the CASLO approach often remained more implicit than explicit. The following analysis is an attempt to explicate them, which has involved an element of retrospective reconstruction based upon the documentary analysis that underpinned this project.

Before explicating these CASLO-specific goals, it is important to distinguish them from goals that were relevant to, but not necessarily premised upon, the CASLO approach. For instance, one of the purposes underpinning the NVQ framework was to provide qualifications that were based on employment-led standards of competence (NCVQ, 1987). This resonated with the ‘new vocationalism’ of the 1970s and 1980s, which emphasised the importance of occupational, in contrast to liberal, objectives. Whereas, previously, attempts had been made to integrate liberal studies within TVET qualifications, NVQs effectively displaced them. While it seems legitimate to argue that this was one of the intended purposes of the NVQ framework, it is harder to argue that this was a motivation for adopting the CASLO approach, *per se*. This is because the spirit of new vocationalism could equally have been operationalised using a quite different approach, even the classical approach.

Note that this spirit of new vocationalism would have been underpinned by even more fundamental sociopolitical goals, such as increasing productivity and international economic competitiveness. These clearly help to explain the perceived need to reform the TVET qualification landscape. But do they explain why reformers turned to outcome-based qualification design, or to the CASLO approach more specifically? Probably not. We need more proximal explanations.

Steve Williams and Peter Raggatt have documented a variety of higher-level sociopolitical goals – goals that are linked to the introduction of competence-based qualifications without necessarily being directly linked to the introduction of the

CASLO approach (Williams & Raggatt, 1996; Williams & Raggatt, 1998; Williams, 1999; Raggatt & Williams, 1999; see also Bates, 1995; Butterfield, 1996). Having said that, some commentators have argued that certain sociopolitical goals were not simply linked to, but specifically relied upon, adopting an outcome-based approach. Indeed, certain sociologists have argued that goals of this sort were paramount in explaining the introduction of outcome-based qualifications. For example:

By defining qualifications in terms of written outcomes alone, an attempt was made to shift the balance of power away from provider-defined qualifications and curricula (which in many instances incorporated professional associations in various ways) towards a broader group of users – government, employers, and learners.

(Young & Allais, 2009, page 8)

We suggest that the emergence of outcomes-based qualifications has been linked to the marketization of education. [...] Learning outcomes or competency statements have come to prominence as a policy tool in this context. They have been seen by policy formulators as a way of driving the required change by playing the role of performance statements in contractual arrangements for educational provision.

(Young & Allais, 2011, pages 2 to 3)

In the first instance, Young and Allais were developing a line of reasoning from Young (2008), which claimed that outcome-based qualifications – based upon employer-driven standards – were introduced because of their potential to help overthrow the educational establishment, constituting: “a completely new approach to vocational education in which (at least in theory) outcomes replaced the curriculum, and workplace assessors replaced teachers” (Young, 2008, page 140). The clarity and transparency provided by learning outcomes and associated assessment criteria was deemed to be critical for empowering those who would assume the roles traditionally played by college teachers.

In the second instance, Young and Allais appeared to be developing a slightly different line of reasoning related to the ability to hold training providers to account. Once again, the idea was that this should help to open training up to new providers who would be held accountable, and compete with each other in a new training market, on the basis of successful delivery of state-endorsed qualifications.

There is certainly some truth in both of these suggestions. We saw how traditional off-the-job college courses were criticised for focusing too much on ‘book knowledge’ and not enough on the ability to perform an occupational role. We also saw how the lack of certification for on-the-job competence was linked to highly variable training provision. NVQs, with their emphasis on workplace learning and assessment were intended to help redress this balance. Moreover, it was true that ministers were keen

to hold education and training providers to account on the basis of assessment and qualification results during the 1980s, and it was assumed that greater use of criterion-referencing might help to promote this. In theory, the more clearly we can define what students are supposed to have learnt during a period of schooling, the more straightforward it will be to hold teachers and trainers to account for whether or not students actually acquire those learning outcomes.

Goals of this sort were clearly important, although whether they were paramount in explaining the introduction of outcome-based qualifications is open to debate.¹²⁷ To reach a conclusion of this sort, we would need to consider the full range of goals that outcome-based qualifications – and the CASLO approach more specifically – were introduced in order to achieve. Therefore, alongside sociopolitical goals, we would need to consider certification and educational goals too.

These 3 categories represent distinct perspectives from which qualification goals can be viewed. The distinction between certification goals and educational ones turns on whether the intention is to improve assessment (certification goals) or to improve learning (educational goals). In the first instance, the intention is to improve the quality of information provided by the certification process, concerning the attainment of each certificated learner. This is often known as improving qualification validity. In the second instance, the intention is to improve the quality of teaching, or the quality of learning, or to increase uptake, or to improve completion rates. These are often described as intended ‘backwash’ impacts from qualification design decisions.¹²⁸

It is clear that, prior to the introduction of outcome-based qualifications, there were longstanding concerns over the validity of existing qualifications – not just TVET ones, but general ones too. In 1943, the Norwood report argued for greater reliance upon teacher assessment, to facilitate a more comprehensive certification process, one that was capable of painting a broader picture of attainment than was possible with external exams. In 1969, the Haslegrave report argued essentially the same point, insisting that traditional external exams failed to test the high-level competencies that technicians needed to acquire, and recommending that more authentic centre-based assessments be adopted instead.

¹²⁷ After all, schools today are held to account on the basis of results derived predominantly from classical qualifications. So, the role of outcome-based qualifications in facilitating the marketisation of education is debateable. Likewise, the idea that outcome-based qualifications were introduced to render colleges redundant as centres of technical and vocational education and training seems hard to square with the fact that they were introduced originally by the TEC and the BEC, agencies that were committed to institutional delivery and to the integration of disciplinary knowledge.

¹²⁸ These distinctions relate to 2 of the 3 perspectives on qualification purposes identified by Newton (2017) – the information perspective and the engagement perspective. The certification perspective is equivalent to the information perspective. The educational perspective is most closely related to the engagement perspective. The sociopolitical perspective does not figure in Newton (2017).

The decision to base TEC and BEC awards upon an outcome-based approach can be seen as a direct response to these concerns, following a logic that had been spelt out decades earlier by Tyler and other pioneers of the Objectives Movement. The outcome-based approach was intended to improve both the comprehensiveness and the authenticity of TVET qualifications – both critical aspects of validity – for TEC awards, BEC awards, and NVQs too.

Following the logic set out by Tyler, outcome-based qualifications were also intended to have positive impacts on teaching and learning. This takes us from certification goals to educational ones. Given their importance in explaining the adoption of the CASLO approach in England, we will explore these educational goals in far more detail below.

Educational goals

Adoption of the CASLO approach in England appears to have been driven by multiple educational goals, which have influenced different qualifications and qualification frameworks in different ways and to differing degrees. The following 4 goals appear to have been particularly important. They relate to improving:

1. domain alignment – to align curriculum, pedagogy, and assessment as closely as possible with the intended domain of learning (and therefore also with each other)
2. domain mastery – to ensure that all students achieve a satisfactory level of attainment across the full domain of learning
3. qualification efficiency – to make the process of becoming qualified as efficient as possible
4. domain personalisation – to enable the domain of learning to be tailored to the personal situation, interests, or needs of learners (or customised to meet the needs of local employers)

These are simply goals, of course, and will not always have been achieved for any particular CASLO qualification or framework. The idea is simply that characteristics of the CASLO approach helped to establish a mechanism by which each of these goals could be achieved. Remember that the 3 core characteristics were defined as follows:

- a. the domain of learning is specified as a comprehensive set of learning outcomes
- b. a standard is specified for each learning outcome, via a set of assessment criteria, and these criteria are used to judge student performances directly
- c. a pass indicates that a student has acquired the full set of learning outcomes specified for the domain

It is important to note that the 4 goals are not achieved through identical mechanisms, which means that they do not attach exactly the same weight to each of the 3 core characteristics. For instance, the first characteristic is particularly relevant to achieving the first goal. This is because close alignment to the target domain requires it to have been specified in full. CASLO qualifications attempt to achieve this via a comprehensive set of learning outcomes, which becomes the single point of reference to which curriculum, pedagogy, and assessment plans can all be aligned.

Of course, even assuming that this facilitates close alignment, we should not assume that learners will naturally end up mastering the domain of learning in full. This would represent a quite different goal, which would require its own mechanism. Of particular relevance to achieving the second goal, domain mastery, is the third characteristic: the stipulation that a qualification will not be awarded unless a student has achieved the full set of learning outcomes.

Although different goals depend on different mechanisms, it is worth noting that one particular mechanism – transparency – is important to achieving all of them. The first and second characteristics, which require learning outcomes and assessment criteria to be articulated, both help to ensure transparency by specifying unit content and standards in detail. So, too, does the mastery requirement, as it makes clear that a qualified individual will have achieved all of the specified learning outcomes.¹²⁹

Before considering how specific CASLO qualifications may have been designed with particular educational goals in mind, we will explain what we mean by each of the 4 goals in a greater detail.

Domain alignment

The domain alignment goal can be traced back to the Objectives Movement, which was pioneered by scholars including Tyler, Bloom, Mager, and others. It seeks to ensure that curriculum, pedagogy, and assessment are aligned as closely as possible with the target domain of learning (and therefore with each other). Misalignment is a classic educational problem, which has been recognised for well over a century, and continues to be a major concern internationally to the present day (see, for example, Porter, et al, 2007; Webb, 2007).

¹²⁹ Transparency is sometimes discussed as though it were a goal in its own right. We think that it is better understood as a requirement, or mechanism, for achieving a variety of goals. For instance, the CASLO approach helps to make it clear to qualification users what the qualification is intended to certificate. This was certainly an important driver behind adopting the CASLO approach. However, we would argue that this is subsumed within the broader goal of making sure that all participants and stakeholders – teachers, trainers, learners, assessors, and users alike – accurately and consistently interpret what the qualification is intended to certificate, which is our domain alignment goal.

The argument for adopting an outcome-based approach to qualification design is that, without a clear and complete specification of a domain of learning, participants in the education-certification process – including teachers, trainers, learners, and assessors – are liable to reach different conclusions concerning exactly what needs to be taught, learnt, and assessed. The idea of a ‘target’ domain suggests that this process needs to be based upon a single, agreed specification of what the qualification is intended to certify and, by implication, what it is not intended to certify. This specification literally becomes the target at which teachers and students aim, and that directs the attention of qualification designers, developers, users, and evaluators alike.

At the heart of the outcome-based approach – and the CASLO approach more specifically – is the idea of specifying the target domain for a qualification as the set of valued learning outcomes that collectively comprises it. The payoff from clear and complete specification is intended to be twofold.

First, it helps to ensure that each qualification targets exactly the right domain of learning. Alternative approaches, including the classical approach, do not specify the intended domain of learning with anything like the same degree of precision. This means that the full scope of the domain may never even be fully debated, let alone agreed.

Second, and more pragmatically, clear and complete specification provides the basis for consistent interpretation. This helps to ensure that the domain of learning is interpreted consistently from one teacher to the next, from one assessor to the next, from one student to the next, from one employer to the next, and so on. It helps to ensure that those in charge of the curriculum interpret the domain of learning in exactly the same way as those in charge of pedagogy and assessment.

If both of these aims are realised, then this paves the way for the target domain to be taught, learnt, and assessed as authentically and comprehensively as possible. This is why domain alignment is simultaneously a certification goal (to improve the quality of assessment) and an educational one (to improve the effectiveness of teaching and learning).¹³⁰

Domain mastery

The domain mastery goal can be traced back to the Mastery Movement, which was pioneered by scholars including Bloom, Carroll, Gagne, and others. It seeks to ensure that all students achieve a satisfactory level of attainment across the full domain of learning.

¹³⁰ Note that we have not identified any CASLO-related certification goals beyond domain alignment.

Domain mastery follows naturally from domain alignment. Having specified the set of valued learning outcomes that collectively comprises a domain of learning, why would we not want all learners to achieve all of them? The CASLO approach involves both a 'stick' and a 'carrot'. The threat of withholding certification constitutes the motivational 'stick' whereas the design of the qualification structure – which helps students and teachers to monitor progress towards completion, with frequent positive reinforcement along the way – constitutes the motivational 'carrot'.

The domain mastery goal holds particular significance for motivating slower learners, who might already have experienced failure on classically designed qualifications, and for whom mastery learning has the potential to be rehabilitative. It also holds particular significance for learners studying occupational qualifications, which are intended to certify full occupational competence. Importantly, though, the domain mastery goal is potentially relevant to any learner in any situation.

Adopting the CASLO approach to achieve this goal means designing a qualification consistent with the expectation that any student who works hard enough ought to be able to acquire all of the specified learning outcomes, to a satisfactory standard, given adequate time and support (and assuming that they began the course from an adequate baseline of knowledge, skill and understanding). It also means assessing each of the specified learning outcomes independently, to be able to confirm when mastery has been achieved. This contrasts with the classical approach to qualification design, which implicitly assumes that only high-achieving students will master the domain of learning in full, and that requires only a relatively small sample of learning outcomes to be assessed. In-course formative assessment is an important requirement for successful mastery learning.

It is sometimes assumed that domain mastery is nothing more than a basic presumption of any occupational qualification. In other words, it is simply of the nature of an occupational qualification that it certifies competence across the full domain of learning. After all, we expect pilots to be able to take off, and land, and do everything flight-related in between. The idea that excellent taking off might somehow compensate for sub-optimal landing does not wash.

It is fair to say that full competence – across each and every specified learning outcome – would be a necessary prerequisite for successfully performing certain occupational roles, although not necessarily all. However, this is much less likely to be true for vocational qualifications taken by school and college students. More to the point, prior to the adoption of the CASLO approach in England, even occupational qualifications were not mastery-based. As such, domain mastery is best understood, primarily, as an educational goal relevant to any qualification context, but of particular significance to certain learners in certain situations.

Qualification efficiency

The qualification efficiency goal is about making the process of becoming qualified as efficient as possible. It applies to the design of qualification frameworks as well as to the design of individual qualifications. Transparency is a critical enabling mechanism, which the CASLO approach achieves by specifying qualification units in terms of both learning outcomes and assessment criteria. Efficiency, in this context, is primarily about avoiding unnecessary duplication of learning and assessment, thereby making qualifications more accessible. It is often discussed in terms of building flexibility into qualifications.

Efficiency is most obviously manifested through qualification systems that accommodate the Recognition of Prior Learning (RPL), which is also known by other names, including the Accreditation of Prior Learning (APL). Specifying qualification units in terms of learning outcomes and assessment criteria enables a learner to determine where they might already have satisfied certain of the requirements for a particular qualification through prior learning. If that prior learning has already been certificated at a sufficiently granular level – which can happen when students pull out of courses having completed some but not all units – their previous achievements can be recognised towards completion of a new qualification. Systems of this sort are particularly useful for individuals who return to learning after a short period of absence, or who move from one location to another and need to transfer from one education or training provider to another. RPL can also be very efficient for those who have not followed a formal course of learning, but who may have acquired competence in situ. Being able to have that informally-acquired competence certificated – against clearly articulated standards – avoids the need to undertake a formal course of education or training unnecessarily (merely for certification).

Efficiency is also a feature of what has become known as Roll-On-Roll-Off (RORO) delivery, which offers candidates the opportunity to begin and complete a qualification at various points throughout the year rather than being restricted to a single point of enrolment and completion. The CASLO approach facilitates this in essentially the same way as for RPL, by tracking the acquisition of learning outcomes (and the completion of units) systematically, synchronously, and transparently.

Efficiency can also be manifested through qualification frameworks that accommodate common units. Specifying units in terms of learning outcomes and assessment criteria enables the determination of identical demands across different qualifications. Units of this sort can be designated as common (to a specified set of qualifications) within the framework. Frameworks of this sort are particularly useful for learners who chose to switch to a different (albeit related) qualification part way through a course of learning. Rather than having to start again from scratch, any

common units will count the same in the new qualification as in the old one. This is sometimes known as Credit Accumulation and Transfer, although it is a strict version of CAT, which depends on exactly the same set of learning outcomes having been specified across qualifications or providers. We might refer to this as a CAT-IA system, as the credits that are transferred relate to Identical Achievements. By making an effort to design common units into qualifications wherever viable, qualification frameworks can be 'rationalised' to avoid unnecessary duplication.

A final manifestation of qualification efficiency, which might be facilitated by adopting the CASLO approach, concerns the potential for recognising progression opportunities. The level of detail provided by learning outcomes and assessment criteria might enable anyone responsible for, say, careers guidance to map plausible progression routes through qualifications with a level of precision that might not be possible if only syllabus content were available for scrutiny.

Domain personalisation

The goal of domain personalisation is to tailor a domain of learning – as set out in the specification for a unit or qualification – to the personal situation, interests, or needs of learners (or to customise it to the needs of local employers). This is often described as an aspect of qualification flexibility. Figure 16 illustrates qualitatively different levels of personalisation. The idea here is that the potential for studying an increasingly bespoke programme of learning increases with each level of analysis.

At the first level, we can imagine the situation for learners who are studying a qualification that comprises only mandatory units, meaning that they all have to achieve exactly the same set of learning outcomes. In this situation, the potential for personalisation would only arise if the learning outcomes were written with sufficient generality to apply across multiple contexts of learning, such that one student might demonstrate them in one context while another might demonstrate them in another.

At the second level, we can imagine a similar situation, but this time with half of the units being optional rather than mandatory. Being able to choose from a variety of optional units means that students are now able to follow somewhat different routes through the same qualification (beyond the common mandatory units).

The third level is intended to reflect the situation in which learners are offered even more potential for customising their programme of learning. This is consistent with the idea of a framework, like the QCF, which offers students the potential to mix and match accredited units to form bespoke qualifications (albeit typically within the parameters of prespecified combination rules).

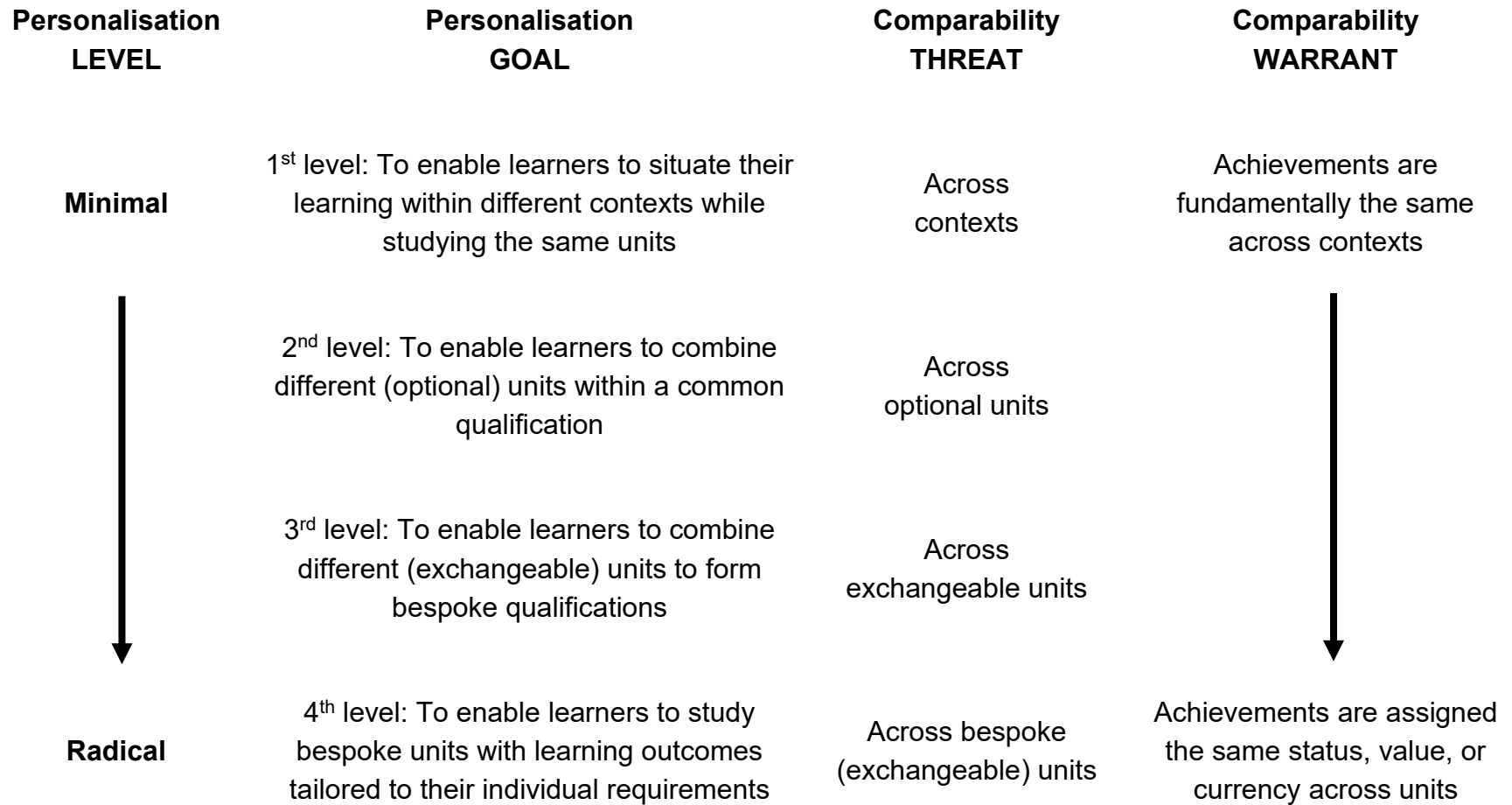


Figure 16. Different levels of personalisation

Finally, the idea of personalisation is taken to an extreme at the fourth level, where unit personalisation becomes possible. This is consistent with the idea of a learner negotiating an entirely bespoke programme of learning, such that even the learning outcomes that comprise each of the units they study can be tailored to their personal situation, interests, or needs.

Figure 16 is intended to be illustrative rather than definitive. For instance, it might be possible to identify mezzanine levels. Alternatively, certain qualifications or frameworks might introduce situations that do not fit neatly into any particular level. Also, it is important to recognise that the levels are not independent of each other, in the sense that units at any level might, for example, incorporate the kind of flexibility illustrated at the first level (although they might not).

The following 2 subsections draw a pragmatic distinction between minimal and radical domain personalisation, linked to the 4 levels from Figure 16. In addition to providing additional exemplification, they explain the significance of the CASLO approach to domain personalisation, including why the first level differs from the remaining ones in this respect.

Minimal domain personalisation

The first level of domain personalisation is the least radical. It corresponds to the situation in which all students within a qualification cohort are expected to acquire exactly the same set of learning outcomes for each unit. The flexibility that is permitted – at this minimal level of personalisation – depends on how tightly the learning outcomes for each unit are circumscribed.

Imagine, for example, that we wanted to design a management qualification to certify the fundamental competencies that managers of small and medium enterprises typically require. Although it might well be possible to reach consensus over a common set of learning outcomes – that might be just as relevant in a small supermarket as in an accountancy firm – it is also quite possible that the expression of those learning outcomes might look quite different across contexts, perhaps even requiring additional ancillary background competencies. An outcome-based approach to qualification design could accommodate a situation like this, as long as the contexts were similar enough for the single set of learning outcomes to have the same significance across contexts, and for the assessment criteria to be equally applicable.

Whenever an element of personalisation is permitted within a qualification, this raises a comparability threat: the greater the degree of personalisation, the greater the threat. In the example just discussed, the issue is whether the different contexts (supermarket versus accountancy firm) are sufficiently similar to be accommodated within a single (management) qualification. This issue has both a logical aspect to it

and an empirical one. From a logical perspective, it would need to be the case that learners really were able to achieve the same outcomes, and to demonstrate the same performance standards, across contexts. Establishing this requirement is an important aspect of qualification design, prior to the qualification going live, which is likely to include debate and consensus building between stakeholders. From an empirical perspective, it would also need to be the case that performance standards were applied, in practice, with the same level of stringency across contexts. This is an aspect of qualification delivery rather than qualification design.

To provide a slightly different example, it might well be possible to construct a set of learning outcomes and assessment criteria – for a particular domain of learning – that had essentially the same significance and applicability for learners who were learning in a college setting as for those who were learning in a workplace setting. The better the college was at simulating a working environment, and the better the workplace was at supporting the acquisition of underpinning knowledge and understanding, the stronger this claim might be. The contexts of acquisition (learning) and demonstration (assessment) would differ, of course, but not necessarily by enough to undermine the claim that students were able to acquire the same outcomes to the same standards.

Although the CASLO approach has the potential to support a minimal level of domain personalisation along these lines, it does not presume it, and can straightforwardly reduce or eliminate it. For instance, outcomes and criteria might be written in a manner that tightly circumscribed the contexts of acquisition and demonstration.¹³¹ Or these contexts might be circumscribed separately, for example, via range statements. Or they might be circumscribed by other qualification design decisions, for example, by controlling assessment tasks more tightly.

By way of summary, the CASLO approach is often associated with the potential for (minimal) domain personalisation because of how explicitly it articulates learning outcomes. Just as detailed specification clarifies exactly what outcomes need to be learnt, it also helps to clarify the degree of flexibility that might be available concerning the contexts within which those outcomes might legitimately be situated, potentially enabling the same outcomes to be acquired and demonstrated in somewhat different contexts. In other words, the transparency of the CASLO approach reveals the potential that exists (or does not exist) for minimal domain personalisation of this sort. This potential can be dialled up or down depending on how tightly the learning outcomes and assessment criteria are circumscribed.

¹³¹ This issue has played out in debates concerning whether the specification of learning outcomes ought to be ‘content-heavy’ (incorporating both syllabus content and cognitive processes in equal measure) or ‘content-light’ (skewed towards cognitive processes with far less prescription over content). Content-light approaches have sometimes been criticised for leading to a “degradation of content” (Hooper, 1971, page 123).

The second level of domain personalisation from Figure 16 is slightly different. It corresponds to the situation in which all students within a qualification cohort study exactly the same mandatory units, but are able to select a certain number of units from a variety of available optional ones. For example, a business qualification might have 4 mandatory units, plus a choice of 2 from the following 4 units: business accounting, financial accounting, management accounting, accounting systems. In this instance, the learning outcomes and assessment criteria would differ across the 4 optional units. Consequently, to justify the provision of optional routes within this qualification, there would need to be an expectation of a reasonable level of comparability of overall demand across the optional units.

Once again, the CASLO approach helps to support this goal via the mechanism of transparency, albeit in a slightly different way. At this level, the detailed specification of outcomes and criteria helps to clarify upfront – at the qualification design stage – the nature, breadth, and depth of the required learning within each of the optional units. If the outcomes within each unit appear to be sufficiently similar in terms of the nature and amount of their content, and if the criteria linked to those outcomes appear to be sufficiently similar in terms of the demands that they make of learners, then this provides a warrant for claiming that the optional units are sufficiently comparable. This is no more than a rough judgement of comparability, of course, based on appearance alone (with no reference to empirical data on how demanding the units actually prove to be in practice). But it is a critical part of the comparability warrant all the same. Incidentally, the transparency provided by the CASLO approach also means that the comparability claim is open to challenge.

Radical domain personalisation

The third level of domain personalisation is more radical than the second because it extends the idea of using (optional) units to construct different routes through a qualification into the idea of using (exchangeable) units to form bespoke qualifications. It corresponds to the situation in which learners are able to mix and match units that have been accredited to a particular location in a qualification framework, guided by rules that specify what sort of units can be combined to form what kind of qualification. As noted above, the QCF provides an example of this approach to domain personalisation.

At this level, the exchangeable units are likely to be less similar in terms of the nature of their learning outcomes. So, the comparability claim that needs to be warranted will be correspondingly looser. Once again, though, it is the detailed specification of outcomes and criteria that helps to clarify upfront – at the unit design stage – the nature (content), breadth (size) and depth (complexity) of the required learning within each unit. This becomes the basis for the judgement that underpins

the accreditation of the unit to a particular location in the qualification framework (characterised in terms of sector area and level, for instance).

This radical domain personalisation goal requires a mechanism for assigning equivalent status, value, or currency to units that certify exchangeable (but not equivalent) achievements. The idea of equivalent status suggests that 2 qualifications might share the same currency in the job market, despite comprising different units and certifying different competencies. According to the logic of this goal, a qualification does not necessarily have to earn its currency over a period of time – by winning the confidence and understanding of stakeholders – if, instead, its currency can be guaranteed upfront by the location that its units occupy within a regulated qualification framework.¹³² For instance, university degrees made up of different modules are commonly accepted as having equivalent status in this sense, where the modules are judged by the university to be (roughly) comparable.

Specifying a common approach to qualification design – particularly a transparent one like the CASLO approach – will help the framework owner to accredit each unit to an appropriate location in the framework. This will be done on the basis of judgements of comparable demand, related to unit content, size and complexity. Occupying a particular location in the framework indicates that a unit has equivalent status, value, or currency to other similarly located units. From this perspective, transparency is the tool that facilitates judgements concerning comparable demand, which provide the justification for claims concerning equivalent value.

The fourth level of domain personalisation is the most radical. It corresponds to the situation in which the content of a unit is tailored to the requirements of an individual learner. This extends the idea of mixing and matching standard units from a central unit bank to the idea developing bespoke units, for individual learners, which can then be assigned currency by accreditation to a location in a qualification framework. Once again, the CASLO approach helps to support this accreditation process by facilitating judgements of (rough) comparability.

These radical personalisation goals are associated with the development of the Credit Movement in England, and with permissive versions of Credit Accumulation and Transfer, which permit non-equivalent achievements to be treated as though they are exchangeable across qualifications or providers. We might refer to this as a CAT-EX system, as the credits that are transferred relate to Exchangeable Achievements (not identical ones). A similar logic underpins the use of an Exemption policy, where entry requirements can be satisfied by a range of exchangeable units, or qualifications, some of which will be more directly relevant to the progression route than others.

¹³² Rather than being earned independently, currency is, in effect, guaranteed by the framework owner, who has accredited the unit to the framework (Wilson, 2010).

Significance of the CASLO approach

It is worth revisiting the significance of the CASLO approach at this point, partly to explain differences between the levels in Figure 16, but also partly to consider alternative dimensions of personalisation and whether they ought also to be described as distinct CASLO goals in their own right. Bear in mind that, by describing something as a CASLO goal, we are suggesting that it counts as a reason why a qualification designer might have opted for (the features that comprise) the CASLO approach as opposed to (the features that comprise) a different approach to qualification design.¹³³

In the case of first level domain personalisation, we might decide to adopt the CASLO approach because the specificity of its outcomes and criteria not only explicate exactly what needs to be learnt, they also explicate the potential for situating that learning in different contexts. Again, for some CASLO qualifications, there may be no such potential, depending on how tightly the outcomes and criteria are specified. Yet, for other CASLO qualifications – where cross-context domain personalisation is an explicit goal – outcomes and criteria will be written specifically to facilitate this.

Although first level domain personalisation raises a threat to comparability, the desire to establish comparability is not the rationale for adopting the CASLO approach. It is, however, the rationale for adopting the CASLO approach for the remaining 3 levels. In each case, the introduction of optional, exchangeable, or bespoke units raises the question of whether these non-equivalent units are sufficiently comparable. If they are not sufficiently comparable, then the qualifications that they result in will also not be comparable, which presents a threat to their currency in the job market. The CASLO approach can help to address this threat by specifying the nature (content), breadth (size) and depth (complexity) of each unit in as much detail as possible. This maximises the potential for reaching defensible conclusions concerning the comparability of non-equivalent units upfront (even before the units have gone live) on the basis of logical analysis alone.¹³⁴

¹³³ The use of parentheses, here, recognises the subtle, but important, point that the CASLO approach is an idea that we have invented (as the conjunction of certain key design characteristics) to make sense of historical policies and practices. For instance, NVQ designers would have adopted the ‘NVQ approach’ which (in our terminology) happens to incorporate the ‘CASLO approach’.

¹³⁴ In theory, it is entirely possible to achieve domain personalisation using classically designed units, for example, by creating optional units to sit alongside mandatory ones within a classical qualification. Yet, the underdetermination of outcomes and criteria within classical unit specifications limits the potential for establishing defensible, albeit rough, comparability claims upfront (that is, on the basis of the specification alone, with no empirical data on the performance of candidates). The more radical

Moving on to alternative dimensions of personalisation, it is sometimes suggested that the following ought also to be described as plausible reasons for wanting to adopt the CASLO approach:

- teaching and learning approach personalisation – which is the flexibility that arises when the acquisition of learning outcomes is compatible with a variety of approaches to teaching and learning (that is, the same outcomes can be taught and learnt in different ways, for example, via didactic instruction or via problem-based inquiry)
- assessment format personalisation – which is the flexibility that arises when the demonstration of the acquisition of learning outcomes is compatible with a variety of approaches to assessment (that is, the same outcomes can be assessed in different ways, for example, via written testing or via oral questioning)

It is true that these are sometimes described as benefits from adopting the CASLO approach. But, should we classify them as CASLO goals in their own right, that is, as substantive reasons for wanting to adopt the CASLO approach (as opposed to a different approach to qualification design)? Or are they better classified simply as fringe benefits from adopting the approach?

It is not obvious, for instance, why the detailed specification of outcomes and criteria – characteristic of the CASLO approach – might make it significantly easier for a teacher to adapt their approach to teaching a particular qualification (compared with a classical qualification that is specified in far less detail). Note that General Qualification examining boards have traditionally shied away from pedagogical prescription, assuming that teachers ought to be free to choose how they teach.¹³⁵

Nor is it obvious why continuous centre-based assessment within a CASLO qualification might be inherently more amendable to assessment format personalisation than continuous centre-based assessment within a classical one. In the CASLO qualification context, assessment format personalisation is often seen as particularly valuable for learners who find responding in a particular format very challenging. For example, a learner might struggle to write a convincing response to a constructed-response exam question, yet they might respond far more convincingly in the context of a professional discussion. Furthermore, in certain CASLO qualification contexts, if a learner fails to perform successfully in one format,

the personalisation, the rougher any comparability claim will be. Yet, it should still be easier to establish (rough) comparability on the basis of a more transparent CASLO unit specification than on the basis of a less transparent classical one.

¹³⁵ It could be argued that the CASLO approach enables learners to take greater control over their learning, making self-study a more viable option. Yet, this potential benefit is already captured by the domain alignment goal, linked to the qualification efficiency goal. Furthermore, it is entirely possible to self-study for a classical qualification, as many learners do.

then this might be directly followed up by assessment in a different format, to see if they are able to perform better.¹³⁶ Yet, whether this is simply a highly valued mitigation within an approach that implements the mastery principle so rigidly, or whether it is a substantive goal in its own right, is open to debate.¹³⁷

Finally, on a slightly different note, it is worth mentioning that the domain personalisation goal can be construed from multiple perspectives. For instance, we have chosen to describe it from a personal perspective, in terms of the benefit that accrues to learners from being able to tailor a domain of learning to their individual situation, interests, or needs (or to the needs of local employers). However, it could also be described from an administrative perspective, in terms of the benefit that accrues to providers from being able to accommodate a wide range of learning needs within a single course structure. This might be considered an administrative goal to the extent that it could have a direct bearing on the viability of course delivery (which resonates with the idea of qualification efficiency).

Qualifications and frameworks

The following subsections identify the goals that appear to have driven qualification designers to have adopted the CASLO approach for key CASLO qualifications from previous sections. We note that design decisions have been driven by different goals for different qualifications.

NVQs and GNVQs

Why was the CASLO approach adopted as the high-level design template for NVQs? From how its designers described and justified the NVQ model, it is possible to identify a number of key drivers.

First, there are strong grounds for believing that the desire to improve domain alignment was the most influential driving force. Remember that the NVQ mission was to develop ‘standards of a new kind’ which were capable of characterising occupational competence comprehensively and authentically, in contrast to the qualifications that preceded them, which were only implicitly defined (in terms of

¹³⁶ There are pros and cons associated with flexibility of this sort, so it needs to be used skilfully. On the one hand, it has the potential to eliminate construct-irrelevant variance (when a learner is unable to demonstrate sufficient proficiency owing to demands associated with a particular assessment format). On the other hand, it also has the potential to create construct-irrelevant variance (when the format provides too much scaffolding, facilitating a successful performance despite the learner being insufficiently proficient).

¹³⁷ Incidentally, we did not find documentary evidence of it appearing as a substantive rationale in its own right as part of the present research project.

syllabus content and exam materials). It was critical that these new qualifications should target exactly the right kind of competence, and that all concerned – from trainers to learners to users – should understand exactly what that competence looked like. This quotation from Jessup could easily have been uttered by Tyler half a century earlier:

The new competence-based movement is attempting to go back to fundamentals and look at what is really required for successful performance or the achievement of successful outcomes in any field of learning. If one faithfully interprets the nature of competent performance, the less tangible skills (including an awareness of context and appropriateness of different responses) will be included as part of the statement. There would appear to be no intrinsic reason why the specification of outcomes should be narrow.

(Jessup, 1991, page 129)

NVQ standards were intended to provide an explicit, common foundation for planning curriculum, for planning pedagogy, and for planning assessment:

From the standards we can derive the curriculum – what the individual needs to learn to achieve the standard – and the assessment system – how the individual will demonstrate that they have achieved the standard.

(Mansfield & Mitchell, 1996, page 85)

Clear and complete explication of agreed standards was intended to improve alignment in contexts such as workplaces where assessments might otherwise have been expected to be highly subjective (Jessup, 1991). Although NVQ designers did not believe that this detailed explication would enable assessors to judge with perfect objectivity, they did believe that the scaffolding provided by outcomes and criteria would facilitate far greater objectivity.

Second, although it made sense that NVQs ought to certify domain mastery – as indicators of occupational competence – Jessup was explicit that mastery ought to be a fundamental educational goal in its own right, and he specifically referenced Bloom on this issue (Jessup, 1991). It is true that NVQ designers were very resistant to prescribing teaching and learning expectations, largely because they wanted to promote the idea of an individual learning journey, especially for those already in work (see below) who were likely to have very different baseline levels of competence and therefore very different learning needs. Yet, Jessup, in particular, was very clear that his new model of education and training presumed that all qualifications, whether technical, vocational, or general ought to be based upon a mastery learning principle.

Third, another very important goal underpinning NVQ design was qualification efficiency. This was perhaps the most explicitly stated goal within NVQ

documentation, although it was articulated in various different ways, often revolving around the concepts of accessibility and flexibility. The transparency afforded by adopting the CASLO approach provided opportunities for individuals to become qualified more efficiently than ever before, for instance, by making use of RPL. This would enable employees to achieve certificates for skills that they had already developed (FEFC, 1997). It even opened the evidence gathering process to sources outside the workplace, including voluntary or leisure activities (NCVQ, 1997c). Enabling already competent, and partially competent, employees to achieve formal certification were key goals for the NVQ system, to promote upskilling and to establish a more mobile workforce (Shackleton & Walsh, 1995). The NVQ model also supported Roll-On-Roll-Off delivery.

The inevitable corollary of RPL was an expectation of tailoring learning experiences to the specific needs of individual learners, rather than forcing all learners to follow exactly the same course of instruction (Jessup, 1991). Talk of common units and credit transfer between qualifications also referenced this goal of qualification efficiency (Jessup, 1991). Finally, NVQ designers believed that the transparency provided by clear and complete standards was educationally empowering, giving learners a degree of control over their own learning that would not have been possible under the classical approach to qualification design. This, too, helped to improve the efficiency of becoming qualified.

Although the NCVQ sought to rationalise TVET qualifications by introducing a qualification framework, the design of this framework was not driven by a strong desire to provide a common currency to support radical domain personalisation via unit exchangeability. The NVQ framework was organised in terms of generic levels, but these indicated little more than a rough hierarchy of occupational roles. NVQs were allocated to a level post hoc, rather than being designed to exhibit a certain level of complexity. Ultimately, each NVQ was tailored to a bespoke occupational role, so its complexity was determined by the demands of the role that it represented.¹³⁸

Turning attention to GNVQs, it would stand to reason that they might share similar goals to NVQs. This certainly seems true in relation to both domain alignment and domain mastery, which appear to have been just as important drivers for GNVQs as for NVQs. However, it is possible to argue that qualification efficiency might have been somewhat less important a driver for GNVQs, which were more likely to have been delivered as sessional courses within mainstream educational settings, with less need for RPL. Having said that, GNVQs still offered the potential for RPL,

¹³⁸ Having said that, formal comparability expectations were included in successive versions of the NVQ criteria, for instance: "It is expected that direct comparability between standards will be achieved between similar or adjacent occupational fields. However, comparison across the entire range of occupational fields will be less exact" (NCVQ, 1988, page 18).

particularly in relation to core skills units.¹³⁹ Furthermore, there was a strong drive within both NVQ and GNVQ traditions toward promoting both student-centred learning and learner autonomy, which definitely was linked to the idea of handing control to learners, to help make the process of becoming qualified more efficient. Likewise, the idea of sharing units across qualifications, and even across qualification types, was beginning to gain traction, and this resonates with the qualification efficiency goal.

Although neither NVQs nor GNVQs appear to have been designed with radical domain personalisation in mind, they do appear to have been designed with (minimal) cross-context domain personalisation in mind, particularly GNVQs. A report from the Further Education Unit put it like this:

Because GNVQs are specified in terms of learning **outcomes** (units of achievement), teachers and learners are able to decide on the kinds of learning activities that will be undertaken in order to achieve the specified outcomes and produce the necessary evidence for assessment.

(FEU, 1994, page 135)

This meant that the acquisition and demonstration of learning outcomes could be tailored to local circumstances, or to students' interests and strengths. Indeed, Ecclestone quoted an NCVQ official who described this as "liberating teachers from the tyranny of curriculum" (Ecclestone, 2002, page 59). Cross-context domain personalisation would also have been relevant to the NVQ model, to the extent that learning outcomes were meant to be specified at a level of generality that would enable them to apply to the same occupational role undertaken with different employers. On the other hand, these occupational roles were tightly defined, and range statements provided further circumscription. So, there was probably less flexibility built into the NVQ model than the GNVQ one.

Finally, it is worth noting that the GNVQ model was promoted during a period that was particularly friendly to the idea of learning styles – which implied the need to tailor learning activities to suit individual learning style preferences – and this also seemed to be accommodated within the emerging GNVQ philosophy (see FEU, 1994). Thus, GNVQs were consistent with the idea of teaching and learning approach personalisation, even if that may not have been a significant part of the underpinning rationale for adopting the CASLO approach.

¹³⁹ Particularly in the college context, RPL could become overly bureaucratic and ultimately inefficient. Also, when granted, it excluded learners from (funding for) relevant learning experiences, which impacted class sizes and course viability.

TEC and BEC awards

Although both the TEC and the BEC designed qualifications with domain mastery in mind, particularly for TEC awards, neither organisation seemed to apply the mastery principle stringently. So, the domain mastery goal was important, but not dominant.

In contrast, both BEC and TEC awards were heavily driven by the domain alignment goal. The qualifications designed by both councils were intended to support rounded programmes of learning, more so than the qualifications that preceded them, which focused on underpinning knowledge and understanding. Their adoption of a precursor to the CASLO approach can therefore be understood best in terms of a desire to specify domains of learning relevant to industry and commerce as comprehensively and authentically as possible, to minimise ambiguity over the ultimate objectives of BEC and TEC programmes. Bear in mind this quotation from the Haslegrave report, which led to the new TEC and BEC awards:

Technicians should be able to extract information from different sources, analyse it and determine the action to be taken, and adjust the action on the basis of its practical effect. The traditional external examination was an unsatisfactory way of testing ability of this kind.

(Haslegrave, 1969, page 53)

The concern, here, was exactly the same as that identified decades earlier by Ralph Tyler: if you fail to specify what learners need to 'do' with the content that they are expected to learn – extract information, analyse it, determine action, adjust action, and so on – then this risks these higher-level functions not being assessed and not being taught. The outcome-based approach adopted by both the TEC and the BEC aimed to mitigate this risk, as both a certification goal and an educational one.¹⁴⁰

Finally, it seems fair to conclude that domain personalisation was also an important driver for TEC and BEC awards. Neither council set out to develop frameworks that were capable of supporting the mix and match approach to qualification design that we have associated with third level radical domain personalisation. Furthermore, although they encouraged colleges to develop bespoke units tailored to the needs of local employers, this was not radical domain personalisation, involving the construction of (fourth level) bespoke units for bespoke qualifications. It would have

¹⁴⁰ It seems reasonable to conclude that BTECs inherited much of their rationale for adopting an outcome-based approach to qualification design from TEC and BEC awards. However, the BTEC route to adopting the CASLO approach was somewhat circuitous, evolving through multiple generations. Indeed, it was only fully adopted during the early 1990s, with mounting pressure to integrate BTECs within the NVQ framework. It is therefore hard to identify a uniquely BTEC-driven rationale for adopting the CASLO approach.

been more like minimal domain personalisation, akin to constructing (second level) optional units for standardised qualifications.¹⁴¹ Note that even their standard units were likely to have been compatible with a certain amount of first level (minimal) cross-context domain personalisation.

OCN awards

The commitment of the National Open College Network to establishing a Credit Accumulation and Transfer system suggests that radical domain personalisation may have been the most influential driving force behind adoption of the CASLO approach for OCN and NOCN awards. Although strongly influenced by NVQ developments, the main reason for adopting the approach for OCN awards was to help determine (and defend) the currency of units within the OCN framework, by clarifying, in terms of learning outcomes and assessment criteria, exactly what credits were being awarded for. Clarity over the content, size and complexity of learning outcomes was intended to enable awards to be assigned to an appropriate location in the proposed new national credit framework. Allocating an award to a particular location would establish its currency, such that awards that occupied the same location in the framework would have the same currency, attesting to their comparability across the Open College Networks (Wilson, 2010). This enabled OCN learners to construct personalised learning programmes with national currency.

To achieve this, rather than representing a rough hierarchy of occupational roles, level descriptors suitable for classifying OCN awards would need to define complexity differently. This was attempted by incorporating ideas from Bloom's Taxonomy (Wilson, 2010). By estimating how long it might take to achieve a set of learning outcomes (size), and by matching those outcomes to a framework level (complexity), the currency of an OCN award within the NOCN framework was established.¹⁴²

The use of the CASLO approach also appears to have been driven by a desire to facilitate domain mastery. That is, with the move towards specification in terms of learning outcomes, the OCNs formally agreed that awards should be contingent upon actually achieving the specified learning outcomes, and not simply upon having completed the programme of study (Wilson, 2010).

¹⁴¹ Lysons (1982) noted that the core and options pattern favoured by the BEC made it more viable to teach groups with divergent course requirements, which describes second level cross-optional-unit domain personalisation from an administrative perspective (rather than a personal one).

¹⁴² As such, adopting the CASLO approach was more a matter of administrative convenience than concern for domain alignment. In other words, the size and complexity of the learning outcomes mattered more than their exact content, which would differ from centre to centre, and potentially from student to student within a centre (Ecclestone, 1992).

The QCF

A final question concerns intentions underlying the design of the Qualifications and Credit Framework. Again, it is not easy to identify these intentions, as they were never set out in terms of the 4 goals described above. Yet, because the QCF appears to have been strongly influenced by the OCN approach, it seems reasonable to conclude that a principal goal influencing adoption of the CASLO approach within the QCF was a desire to facilitate radical domain personalisation – albeit at the third rather than the fourth level – implying that the CASLO approach was adopted to help support claims concerning comparable demands, common currency, and exchangeability.

It has certainly been said that parity within the QCF provided a basis for claiming that qualifications should be treated equivalently within accountability measures, for example, underpinning the idea of a ‘GCSE equivalent’ qualification for key stage 4 performance table computations (Wolf, 2011). Furthermore, a guidance note on ‘Writing QCF Units: How Much Detail to Provide’ (FAB & JCQ, 2010) was very explicit on how the learning outcomes and assessment criteria within QCF units needed to be written: to provide the level of detail necessary to indicate the ‘amount’ of learning required (to allocate credit), to indicate the ‘demand’ of the learning required (to allocate levels), and to indicate the ‘content’ of the learning required (to ensure that the units would refer to equivalent achievements when employed by different awarding organisations). This was intended to support comparability judgements related to accreditation decisions, and ultimately to support credit transfer across providers.¹⁴³

In fact, the transparency that was necessary for qualification efficiency CAT (CAT-IA) was probably more important than the transparency that was necessary for radical domain personalisation CAT (CAT-EX). In practice, the mix and match functionality of the QCF was not used very much. Indeed, QCF rules on combining units actively prevented qualifications being assembled “in real time” by employers or learners, which actively constrained radical domain personalisation (Lester, 2011, page 210).

There is little clear-cut evidence from early QCF documentation that either domain alignment or domain mastery were especially strong influences. However, it is quite possible that these goals might simply have been taken for granted, given how embedded the NVQ approach had become by then.

The analysis of CASLO qualifications and frameworks in terms of their goals helps to remind us that the CASLO approach is nothing more than an approach (that is, one approach) to achieving domain alignment, domain mastery, qualification efficiency,

¹⁴³ It noted that it would generally not be possible to port (less transparent) NQF specifications directly into the QCF.

and domain personalisation. It is entirely possible to envisage alternative approaches to qualification design that are capable of achieving these high-level goals in different ways. Indeed, there might conceivably be better approaches to facilitating them – better than the classical approach and better than the CASLO approach.

This observation is particularly relevant to our discussion of the QCF. As noted earlier, the QCF was designed on the assumption that it would ultimately encompass all formally assessed learners' achievements outside higher education. Yet, prioritising qualification efficiency and radical domain personalisation resulted in 2 related problems. First, the QCF ended up being designed to meet the needs of only a subset of learners (and awarding organisations) for whom these goals were clearly relevant.

Second, because these goals recommended a common approach to qualification design across the framework – for which the CASLO approach was chosen – it meant that qualifications that were not well suited to this common approach were distorted. As we have already discussed, this included graded performance exams, including graded exams in music, dance, speech, and drama. It is worth emphasising that these qualifications were already based on a mastery model – the progressive mastery model – and they had their own approach to securing domain alignment. More to the point, they had their own well-established framework, and learners did not stand to benefit from the kind of Credit Accumulation and Transfer anticipated by the QCF. In short, 2 of the QCF design goals were of limited or no relevance to graded performance exams, while the other 2 were already achieved via alternative approaches. In retrospect, it seems fair to conclude that the QCF was designed in a manner that was unsuitable for regulating the full range of qualifications that were eventually accredited to it.¹⁴⁴

¹⁴⁴ Note that the 'unsuitability' case was accepted for certain qualifications (including GCSEs and A levels, yet not for others (including graded performance exams).

Chapter 6. Recent history

Regulatory arrangements were published in August 2008 to support implementation of the QCF (Ofqual, 2008a). This document specified that all QCF units had to be written entirely in terms of learning outcomes and assessment criteria, and indicated that passing a unit meant mastering all specified learning outcomes. The dominance of the CASLO approach became evident as the vast majority of regulated qualifications transitioned into the QCF (typically, to qualify for public funding).

Just a few years later, however, both the QCF and the CASLO approach were to be called into question following a series of official policy reviews. In 2015, the QCF was withdrawn, meaning that there was no longer a regulatory requirement for any qualification in England to adopt the approach. Indeed, an increasing number of qualifications and assessments were prohibited from doing so (a trend that started even before the QCF had been withdrawn).

Post-2010 policy reviews

After the May 2010 general election failed to return a single governing party, the Conservatives and Liberal Democrats formed a coalition. Although plans for reforming Technical and Vocational Education and Training did not loom large in their initial 'Programme for Government' (HM Government, 2010), a series of high profile policy reviews subsequently paved the way to substantial reforms, which continued following the next election when the Conservatives won a majority.

None of these reviews focused specifically on the CASLO approach to qualification design. However, the approach did feature in many of them, sometimes obliquely and sometimes more directly. Whereas, prior to 2010, official reviews tended – either implicitly or explicitly – to support the approach, the post-2010 reviews tended to be more critical. We will consider how the approach featured within 5 of these reviews, and how this influenced subsequent policies and practices related to the CASLO approach:

- Wolf review – 14 to 19 vocational education (Wolf, 2011)
- Richard review – apprenticeships (Richard, 2012)
- CAVTL review – adult vocational teaching and learning (CAVTL, 2013)
- Whitehead review – adult vocational qualifications (Whitehead, 2013)
- Sainsbury review – technical education (Sainsbury, 2016)

Wolf report

In his Foreword to the Wolf report, Secretary of State for Education, Michael Gove, explained that he had invited Alison Wolf to confront a longstanding “failure to provide young people with a proper technical and practical education of a kind that other nations can boast” (Wolf, 2011, page 4). Wolf approached this challenge with a wealth of experience, which – from the perspective of the present report – included having conducted substantial research and analysis into NVQs and GNVQs, and having written a landmark book on Competence-Based Assessment (Wolf, 1995).

Wolf was critical of the state of vocational education in England, concluding that many students were being let down. She was particularly critical of the plethora of low-level qualifications that students were encouraged to take, which had little or no labour market value. This meant that many 14 to 19-year-olds left education without the skills that would enable them to progress.

Recommendations from the Wolf report were extremely wide ranging, addressing a host of issues related to curriculum, teaching, qualifications, apprenticeships, funding, work experience, employer involvement, regulation, accountability measures, and others too. At the heart of these recommendations was the intention that all 14 to 19 students should be following a valuable programme of learning with clear progression opportunities, and that all supporting systems and structures – including qualifications, funding, accountability, regulation, and so on – should clearly align with this intention.

Wolf foregrounded the problem of perverse incentives related to funding and accountability, which steered learners away from coherent programmes and encouraged them to focus upon accumulating easy-option, low-value qualifications. Indeed, she noted the trend for students to be channelled away from high-value academic qualifications toward low-value vocational ones, which were easier to achieve despite attracting equivalent school performance table points. This helped schools to (appear to) perform well but ultimately let students down.

Wolf argued that 14 to 19 learners should not be studying highly-specific qualifications that were unsuitable for them, including those designed specifically for adults working towards occupational competence, and based solely on National Occupational Standards. Because young people change sectors, occupations, and jobs very frequently during their first decade of employment, they needed to study “fairly general” vocational qualifications (Wolf, 2011, page 74). Schools and colleges should therefore be incentivised to ensure that vocational qualifications studied by 14 to 19 students were suitable and valuable:

Only those qualifications – both vocational and academic – that meet stringent quality criteria should form part of the performance management regime for schools.

(Wolf, 2011, page 11)

Government was to put this principle at the heart of its response to the report.

Concern over the QCF

Although Wolf did not focus specifically on the CASLO approach, she did have plenty to say about the (newly established) Qualifications and Credit Framework, which stipulated the approach. She also had a lot to say about QCF qualifications that were offered in schools, particularly those had been allocated the same performance table points as GCSEs. Nominal equivalence was not the same as substantive equivalence, Wolf insisted, and vocational students who took ‘GCSE equivalent’ QCF qualifications had been sold short. Furthermore, she argued, QCF qualifications – designed for adults with narrowly defined occupational goals – should not be the main, let alone the only, type of vocational qualifications offered to students in schools and colleges.

The report also flagged up various design features associated with QCF qualifications related to their association with National Occupational Standards. Of particular relevance to our discussion of the CASLO approach, she proposed that the following characteristics make QCF qualifications ill-suited to delivery within education and training institutions:

- as QCF qualifications require students to achieve all specified learning outcomes, this mastery requirement means that “no single element can be difficult” (page 88) because, if a student fails one element, then they fail the qualification
- the same requirement – with its heightened risk of students failing – also places “enormous downward pressure on standards” (page 87) in the context of teachers assuming considerable responsibility for assessing QCF qualifications, given the perverse incentive not to fail students when institutions are paid by results
- the mastery requirement also incurs “large costs in time and money spent assessing, recording, re-assessing, etc.” (page 88)

Recognising problems such as these, Wolf proposed that:

If awards are to be used for national performance monitoring, it is vitally important that there be very strong safeguards against downward pressure on standards. It would be nice to think this is unnecessary, but the experience of the last few years tells us otherwise. All those which are used, vocational or academic, should make serious demands of students, develop and accredit distinctive skills and attainments, facilitate progression post-16 and incorporate *clearly established, and properly monitored, national standards*. They must, therefore, have a strong element of external assessment. *This need not, and indeed should not, mean*

assessment entirely on the basis of examinations, which in the case of vocational awards will often be quite inappropriate. But we know that, without regular external referencing, assessment standards in any subject invariably diverge across institutions and assessors.

(Wolf, 2011, page 112)

Government's Response

Government accepted recommendations from the Wolf report without reservation, and proposed to implement them in both letter and spirit. Its action plan identified 3 major themes, the second of which promised to:

Reform performance tables and funding rules to remove the perverse incentives which have served only to devalue vocational education, while pushing young people into qualification routes that do not allow them to move into work or further learning. Those vocational qualifications that attract performance points will be the very best for young people – in terms of their content, assessment and progression.

(DfE, 2011a, page 3)

This would be achieved by tightening the accountability system to ensure that only certain vocational qualifications would be counted in school and college performance tables – only ones that government deemed to be respected, and comparable to others in the tables in terms of the rigour of their content and assessment. These qualifications would need to meet stringent quality criteria, meaning that awarding organisations would need to undertake a considerable redevelopment programme.

Performance table requirements

The Department for Education (DfE) developed its response in 2 phases. It launched a consultation on qualifications and performance tables for 14 to 16-year-olds immediately (DfE, 2011b), followed 2 years later by consultations on vocational qualifications for 16 to 19-year-olds (DfE, 2013a) and performance tables for 16 to 19-year-olds (DfE, 2013b). Alongside the first consultation, the Department published technical guidance for awarding organisations (DfE, 2011c), which provided further insights into its proposals and their rationales. This began by stating that:

In line with Professor Wolf's recommendations, in future, only qualifications which are high quality, rigorous and enable progression to a range of study and employment opportunities will be recognised in school performance tables for 14-16 year olds.

(DfE, 2011c, page 1)

It went on to insist that vocational qualifications “must be just as stretching and challenging as academic or general qualifications” (DfE, 2011c, page 1). To demonstrate this, all performance table qualifications would need to satisfy DfE criteria related to size, grading, external assessment, synoptic assessment, progression opportunities, proven track record, and appropriate content (which would be determined, up front, via a formal evaluation process). These rules included a minimum of 20% external assessment: “to ensure that vocational qualifications offer a comparable level of challenge to academic qualifications and are seen to do so” (DfE, 2011c, page 6). They also included an unspecified amount of synoptic assessment:

Synoptic assessment is vital to increase the level of challenge for students as it requires a broader comprehension. This will help ensure that vocational qualifications are as challenging as academic ones. Taken with the minimum size requirement, adding synoptic assessment will ensure cohesiveness across a qualification and prevent qualifications from being treated as a series of disconnected components.

(DfE, 2011c, page 7)

Note that the synoptic assessment requirement responded to a frequent criticism of CASLO qualifications, and unitised qualifications more generally, which is that they can lead to fragmented teaching and learning. Although Wolf would have recognised this criticism, it was not actually foregrounded in her report. External assessment was mentioned, of course, although more to secure confidence in the application of national standards than to secure parity of standards between vocational and academic qualifications.

Technical requirements for school and college performance tables were refined over time. By 2017, requirements for Technical Awards (14 to 16) and for Technical Certificates, Technical Levels, and Applied Generals (16 to 19) were as summarised in Table 6 (see DfE, 2017).¹⁴⁵

¹⁴⁵ These categories were defined as follows (see DfE, 2017, page 5):

Technical Awards – high quality Level 1 and 2 qualifications that equip 14 to 16 year olds with applied knowledge and practical skills.

Technical Certificates and Tech Levels – Level 2 and 3 qualifications that equip post-16 students with the knowledge and skills they need for skilled employment or for further technical study.

Applied General qualifications – Level 3 qualifications for post-16 students who wish to continue their education through applied learning.

	14 to 16 Qualifications	16 to 19 Qualifications
Declared purpose	required	required
Size	at least 120 GLH	at least 150 GLH (AG) at least 150 GLH (TC) at least 300 GLH (TL)
Employer or HE recognition	not applicable	required
Appropriate content	“set out in the specification or supporting documentation clear information about the content of the qualification; this should be more than just the learning outcomes and assessment criteria”	at least 60% mandatory (AG) at least 40% mandatory (TC) at least 40% mandatory (TL)
Appropriate assessment	at least 40% external internal assessment verified or moderated suitably controlled	at least 40% external (AG) at least 25% external (TC) at least 30% external (TL) internal assessment verified or moderated suitably controlled
Synoptic assessment	‘sufficient’ synopticity	‘sufficient’ synopticity
Grading	Pass, Merit, Distinction or more detailed	Pass, Merit, Distinction or more detailed
Employer involvement	not applicable	for TC and TL only
Progression opportunities	required	required
Track record	required	required

Table 6. Technical guidance for awarding organisations.

These DfE requirements effectively ruled out adopting the CASLO approach, at least at the qualification level, given the likelihood of externally assessed units adopting a classical approach (based upon numerical marking as opposed to direct judgement

against assessment criteria). Having said that, it would still have been possible for awarding organisations to develop hybrid qualifications, with internally assessed units adopting the CASLO approach and externally assessed units adopting a classical one. Note that the ‘appropriate assessment’ criterion permitted ‘verification’ which tended to be associated with quality assurance within CASLO qualifications. Many Level 3 BTECs, for instance, were hybridised through the reform process.

Richard report

In June 2012, the entrepreneur, Doug Richard, was asked by the Secretary of State for Education (Michael Gove) and the Secretary of State for Business, Innovation and Skills (Vince Cable) to consider the future of apprenticeships in England, and to recommend how they can meet the needs of a changing economy.

Richard was radical in his response, arguing that apprenticeship should be redefined – with the relationship between apprentice and employer at its heart – as a high skill, high status pathway. Apprenticeship should no longer be seen as a government-led training scheme, dominated by training professionals. It should be understood as an employer-led educational journey for an apprentice who is new to their role and has much to learn. In exactly the same way, standard setting and assessment should also be employer-led, rather than being dominated by Sector Skills Councils and awarding organisations.

Richard raised concerns of direct relevance to the CASLO approach, albeit couched within broader concerns over apprenticeship standards and assessment.

Concern over standards

Central to the Richard report was the idea that apprenticeships no longer provided a guarantee of overall occupational competence. This was primarily due to apprenticeship frameworks in which a “welter of qualifications” acted “like stepping stones” yet often without ever declaring apprentices competent (Richard, 2012, page 4). Worse still “we have an extraordinary number of qualifications, which under the guise of flexibility can be stitched together in an infinite number of combinations leading to any possible outcome but no clear accomplishment” (Richard, 2012, page 6).

This lack of coherence across qualifications was compounded by micro-level specification of National Occupational Standards, which made it hard to see the wood (overall competence) for trees (elements of competence). Worse still, this inadvertently constrained innovation and flexibility in teaching, and meant that apprentices spent too much time being assessed and not enough time being trained:

Too much provision is however driven by the need to tick off a very long list of competencies, required to complete the requisite qualifications. This has meant that, today, too many apprenticeships involve, in part if not [in] total, a heavy focus on on-going assessment – indeed many apprenticeships are delivered on the ground almost exclusively by individuals called assessors, rather than trainers, teachers or educators. Much of the time which apprentices spend ‘training’, is in fact spent with their assessor providing evidence of their ability to meet competency requirements. I believe apprenticeships should be about new learning, and those involved in delivering apprenticeships should focus on teaching and coaching – this should be their primary task, the thing they are paid to do.

(Richard, 2012, page 87)

Streamlining would provide a solution to these problems. There should be just a single qualification for each apprenticeship, and its outcome-focused standard:

should clearly set out what apprentices should know, and be able to do, *at the end* of their apprenticeship, at a high level which is meaningful and relevant for employers

(Richard, 2012, page 17)

The effect would be to simplify the system, freeing up curriculum and pedagogy at the same time. Quality, in this newly simplified system, would be underpinned by strong leadership from employers.

Concern over assessment

The idea that apprenticeships failed to provide a guarantee of overall occupational competence also influenced recommendations concerning assessment, which also risked losing sight of the wood for the trees:

Finally, we know that success in an individual qualification or component of an apprenticeship does not always guarantee competence in actually doing the job. Employers tell me that individuals could tick off the many tasks involved but not, at the end, be genuinely employable and fully competent.

(Richard, 2012, page 50)

Streamlined standards would result in streamlined assessments, which would help to solve this problem too. These streamlined assessments – scheduled for the end of an apprenticeship over a period of days or weeks – would be far more integrated:

The final test and validation must be holistic, in that it seeks to test the full breadth of the relevant competencies not merely the incremental progression of the apprentice. That may take the form of a project or an assessment in front of

an examiner. It should be performance and real world based, rather than just theoretical. It should be primarily at the end of an apprenticeship, not measuring progress during it.

(Richard, 2012, page 8)

Quality, in this simplified assessment system, would be underpinned by strong leadership from employers. However, Richard believed that there was also a need to externalise the system, to underpin its credibility. This was not in the sense of insisting upon external written exams. Indeed, Richard proposed that the test “will need to be primarily practical and involve directly observing whether the apprentice can do their job well, in different and novel circumstances” (Richard, 2012, page 54).¹⁴⁶ Instead, externality would be provided by appointing assessors who were independent of anyone with a strong interest in the apprentice passing.

Government’s response

Matthew Hancock, Minister of State for Skills, received the Richard report on behalf of the DfE and BIS, and set out the government’s next steps for consultation. He concluded that the report set out a compelling case for reform, which would ensure that apprenticeships become “rigorous and responsive” to employer needs (DfE & BIS, 2013, page 3). Post-consultation decisions were set out in ‘The Future of Apprenticeships in England: Implementation Plan’ (HM Government, 2013a).

On standards, government decided that:

In future, Apprenticeships will be based on standards designed by employers to meet their needs, the needs of their sector and the economy more widely. These standards, which will replace the current frameworks, will be short, easy to understand documents that describe the level of skill, knowledge and competency required to achieve mastery of a specific occupation and to operate confidently in the sector.

(HM Government, 2013a, page 4)

The new standards were therefore intended to embrace both practical and theoretical elements.¹⁴⁷

On assessment, government decided that:

An apprentice will need to demonstrate their competence through rigorous independent assessment, focused primarily on testing their competence at the

¹⁴⁶ Although, he acknowledged that the test might also assess knowledge and understanding if required by the industry in question.

¹⁴⁷ Guidance from 2014 onwards would refer to 3 core constructs – skills, knowledge, and behaviours.

end of their Apprenticeship. The assessment will be against the relevant standard, and employers will have a key role in developing the high level assessment approach.

(HM Government, 2013a, page 4)

Grading was also to be introduced, to encourage apprentices to strive for excellence. Of relevance to the CASLO approach, the new standards were still required to be outcome-based, and they were still required to certify mastery. However, the outcomes were to be defined very much more succinctly, and mastery was to be understood correspondingly holistically:

The new standards will be short (typically one side of A4), easy to understand documents that describe the level of skill, knowledge and competency required to undertake a specific occupation well, and to operate confidently within a sector. They will focus on how an apprentice should demonstrate mastery of an occupation, and will not list narrowly defined tasks.

(HM Government, 2013a, page 11)

Indeed, on the same page, the report appeared to redefine 'mastery' in terms of the need for an apprentice to be able to transfer their skills when moving from one company to the next in the same occupation. This was consistent with the principle of focusing each standard upon a broadly defined occupational role rather than a narrowly defined job (with a particular employer, which might only require a subset of the skills required for essentially the same role with another employer).

Having said that, just a few pages later, the report explained that apprentices would still have to demonstrate "their ability in all areas of the standard" (HM Government, 2013a, page 15), which suggested that mastery still meant jumping a series of hurdles. Indeed, this was clearly spelt out:

Grading will be applied to the full Apprenticeship standard and a mastery mechanism of assessment will be used. This means that apprentices will need to pass every aspect of their assessment in order to be successful, but not every aspect will need to be graded. This approach means that an apprentice will not be able to compensate for failure in any one aspect of the assessment with a strong performance in another area.

(HM Government, 2013a, page 18)

Finally, reflecting the idea of holistic competence, the document explained that assessment would include:

a synoptic element to the end-point assessment, requiring the apprentice to identify and use effectively in an integrated way an appropriate selection of skills, techniques, concepts, theories, and knowledge from across their training

(HM Government, 2013a, page 17)

Both standards and assessment plans were to be developed by ‘Trailblazer’ groups, comprising leading employers and professional bodies. An accompanying document ‘Guidance for Trailblazers’ (HM Government, 2013b) explained these decisions in slightly more detail. The desire to put employers in the driving seat meant that they should have as much freedom as possible when developing standards and assessment plans. This included the option of mandating the achievement of existing qualifications within the standard, if they wished to.¹⁴⁸ Where rules were specified, they tended to be fairly loose, such as the requirement for the specification of a standard to be “concise (typically around one side of A4)” and “written in clear and simple language” (HM Government, 2013b, page 14). The concept of ‘mastery’ was not elaborated in this guidance.

Subsequent guidance extended the anticipated length of each standard to “one to two sides” (HM Government, 2014, page 23).¹⁴⁹ It also seemed to loosen the mastery requirement somewhat, given the pragmatic need to sample that arises when assessment is no longer continuous:

The end-point assessment must assess across the whole standard but it does not have to assess every aspect. When thinking about which aspects of the standard would need to be formally assessed at the end of the programme, it may be helpful to think about how critical it is for the occupation, how frequently it is used and whether it links to professional registration.

(HM Government, 2014, page 44)

Yet, the passing grade was still intended to certify full competence, and this expectation continued across subsequent iterations of the guidance.

The next iteration provided further elaboration of what a successful apprenticeship standard might look like:

At the core of a successful apprenticeship standard are two things:

- A short and clear role description setting out the main activities that someone in this occupation would do, in language that can be easily understood by someone without technical knowledge.

¹⁴⁸ Indeed, some mandated the achievement of an NVQ.

¹⁴⁹ This stipulation related only to the standard. The assessment plans would have been longer.

- A definitive list of the skills, knowledge and behaviour that you as an employer would expect from someone who is a fully competent professional in the occupation.

(HM Government, 2015, page 17)

Trailblazers were still restricted to 2 sides of A4 (size 12 font), unless proposing a ‘core and options’ approach, which permitted slightly more space.

Decisions on the design of apprenticeship standards and assessment plans, which stemmed from recommendations in the Richard report, represented a shift away from the CASLO approach. Yet, how radical a shift this was to be remained a little unclear. On the one hand, Richard bemoaned the kind of micro-level specification that had been associated with NOS and NVQs. He argued that continuous assessment of finely specified standards meant that apprentices were committing too much time to assessment – time that would be better spent on training. The switch to external end-point assessment was intended to render it shorter, more holistic, and no longer reducible to an exercise in box ticking. On the other hand, Richard still wanted apprenticeships to be defined in terms of outcomes, and still wanted the passing grade to certify full occupational competence. This translated into guidance that sometimes sounded very reminiscent of the CASLO approach, which left open the possibility that assessment under the new approach might still be reduced to an exercise in box ticking, albeit with fewer boxes to tick (bearing in mind how streamlined the new standards had become).

The Federation for Industry Sector Skills and Standards (FISSS) managed apprenticeship certification under the framework system, and was to continue doing so under the new standards system.¹⁵⁰ During 2014 and 2015 it published a series of reports intended to help Trailblazer groups, and ‘enablers’ of those groups (including professional bodies and Sector Skills Councils) to respond to the 2013 ‘Implementation Plan’. Its ‘toolkit’ for enablers (FISSS, 2014) provides interesting insight into the variety of ways in which the first Trailblazer groups approached their remit. For instance, although it stated that there “may be more value in starting afresh” it recognized that many of the first Trailblazers “based their respective standards primarily on the existing framework, as it already met employer requirements” (FISSS, 2014, page 21). It also continued to endorse functional analysis, with a nod to the skills built up by its members:

- Functional elements: How are the key functional elements of the occupational competence – professional skills, knowledge, and (optionally) behaviours – identified and agreed? *Functional analysis is the recognised approach.*

¹⁵⁰ In 2008, the Sector Skills Development Agency was [replaced](#) by the UKCES and the Federation for Industry Sector Skills and Standards comprising all 19 sector skills councils.

- Does the group need support with the functional analysis of the occupation or job role to draw out the relevant skills, knowledge, and behaviour that demonstrate competency? *Professional bodies and sector skills councils have extensive technical knowledge of standard development.*

(FISSS, 2014, page 21)

CAVTL (McLoughlin) report

In December of 2011, the Department for Business, Industry and Skills (BIS) set out plans for reforming the further education and skills system, which included actions to develop and promote excellent teaching (BIS, 2012). An independent commission on adult education and vocational pedagogy would be established with a remit to develop a sector-owned strategy and delivery programme.

The Principal of City & Islington College, Sir Frank McLoughlin, chaired this Commission, which reported in March 2013. Reflecting an intention to speak on behalf of a range of stakeholders – including industry stakeholders, teaching trainers and practitioners, and professional associations – their report tends to be known as the CAVTL report (Commission on Adult Vocational Teaching and Learning). It focused on 18+ learners on vocational (but not pre-vocational) programmes.

The report was titled 'It's about work' and stressed that learners require a clear line of sight to work to be able to appreciate exactly why they are learning what they are being asked to learn (CAVTL, 2013). At the heart of the report was the idea of strengthening links between teaching and learning (on the one hand) and employers and employment (on the other). This was embodied in the idea of creating a two-way street, which meant that employers should be full partners in the further education and skills system (alongside trade unions and professional bodies) and not mere customers to colleges and training providers.

Genuine collaboration was key to establishing this two-way street. Employer involvement and influence would need to be improved, including direct involvement in curriculum planning. In turn, this would require more flexible qualifications, which could be tailored to local needs. The Commission recommended a 'core and tailored' approach, that is, a national core plus a tailored element to meet local demands.

The vision of a two-way street would need to be supported by excellent teachers and trainers – dual professionals who combined occupational expertise with pedagogical expertise. But these professionals would need training and development, which would require substantial investment. The need to invest in professional updating was identified as a particular priority.

In addition to multiple site visits, McLoughlin commissioned a series of briefing papers. The first, provided by Geoff Stanton, former Director of the Further

Education Unit, was discussed at the first meeting of the Commissioners. It is worth mentioning because of its discussion of vocational qualifications. Stanton emphasised how effective teaching and learning depends on striking the right balance between qualification design, the development of learning programmes, and pedagogical planning. He noted that, at different points in time, these different factors had been weighted differently. With the introduction of NVQs, the emphasis now lay very heavily on qualification design. A once iterative process had now become extremely linear:

- first develop occupational standards
- then develop qualification structures and processes to suit these standards
- then develop learning programmes to suit these structures and processes

Unfortunately, according to Stanton, this model had resulted in some very negative consequences for vocational pedagogy. Because teachers were not involved in the specification of standards:

some outcomes though measurable proved very difficult to teach, some important learning experiences were neglected because they could not be easily reflected in fundable outcomes, and many of the standards were expressed in terms that were incomprehensible to those hoping to achieve them

(Stanton, 2012, page 7)

Although he recognised that the advent of NOS had enabled much previously unstructured training to be systematised, which was important, he also noted the tendency for trainers to use those standards directly as the basis for qualification delivery – as though they constituted a learning programme – rather than indirectly as the foundation for course development. He implied that we should not be too surprised about this, bearing in mind that, while trainers were required to possess a qualification in NVQ assessment, they were not required to possess a qualification in NVQ teaching and learning.

The Commission echoed these concerns, particularly those related to the tendency to construct learning programmes directly upon NVQ-mediated NOS, with little if any attention paid to curriculum progression and pedagogical implications:

We need to put curriculum development and programme design back at the heart of vocational teaching and learning. Over the last 30 years, the emphasis has shifted from curriculum development to qualifications design, which has wrongly been equated with programme design. Together with a funding regime based on qualifications, this has exacerbated a focus on '*assessment as learning*' and qualifications.

(CAVTL, 2013, page 14)

The newly formed Education and Training Foundation was charged with taking forward recommendations from the CAVTL report, although the Commission envisaged that certain of the recommendations – including its recommendation concerning ‘core and tailored’ qualifications – would be developed in the forthcoming review of adult vocational qualifications.

Whitehead report

In [spring 2013](#), Matthew Hancock asked Nigel Whitehead, BAE Systems Group Managing Director and Commissioner for the UK Commission for Employment and Skills (UKCES), to review adult vocational qualifications in England. As discussed earlier, the UKCES inherited responsibility for managing National Occupational Standards when it was established in 2008, soon committing to a substantial reform programme, which was rolled out in 2011. Recall that the Richard review had criticised these reformed NOS, and Matthew Hancock formally responded to these concerns in spring 2013 (DfE & BIS, 2013). The new model of apprenticeship standards was confirmed a month prior to the Whitehead report being published (see HM Government, 2013a).

Whitehead interpreted his remit in terms of seeking out issues for improvement and providing recommendations for reform. He identified a number of “systemic weaknesses and unintended outcomes” (Whitehead, 2013, page 3) and presented a vision for reform designed to ensure that adult vocational qualifications would become more:

1. relevant – being linked directly to occupations, either to support entry into an occupation, or to provide professional development within it
2. rigorous – being more reliable, robust, and graded, derived from clear and ambitious occupational standards, and not constrained by QCF design rules
3. recognised – being better understood and respected, with better data on progression and returns

Underpinning this vision was the principle of employer ownership. Adult vocational qualifications should be driven by business leadership rather than by government management. Employers should take end-to-end responsibility for workforce development, working in partnership with competitors, supply chains, unions, training providers, professional bodies and awarding organisations.

Whitehead interpreted ‘vocational’ to mean qualifications linked directly to occupations. He excluded from this category – and from his review – low-level, confidence-building qualifications designed to recognise progress towards the labour market. Indeed, he saw the lack of identity of adult vocational qualifications as a problem in its own right, recommending that Ofqual should regulate them as a qualification type with their own design principles. Making them more relevant to

employment would also help to improve their identity, giving them a clear line of sight to a job or to a range of jobs, consistent with recommendations from the CAVTL report, and reflecting concerns from the Wolf report over the prevalence of low-value qualifications.

Whitehead recognised numerous observations and recommendations from Wolf, Richard, and CAVTL. The following 3 sections illustrate this, while discussing issues of particular relevance to the CASLO approach.

Standards

Whitehead echoed Wolf, Richard, and CAVTL in noting that over-detailed NOS risked constraining both teaching and assessment. He agreed with Richard that we should move away from NOS written with excess detail toward “clear high level outcome-based” standards (Whitehead, 2013, page 31). He applied the same reasoning to QCF qualifications, reiterating the concern expressed by CAVTL that detailed specification of assessment criteria risked creating a culture of ‘assessment as learning’. Whitehead was particularly concerned that overly prescriptive standards constrained training providers, making it difficult for them to customise the curriculum to meet local needs.

More specifically, he recommended that adult vocational qualifications should satisfy a set of design principles, one of which was that they should allow for a proportion of locally-specified standards. He described this as a ‘core and options’ model. This would enable qualifications to meet the needs of specific industries, small- and medium-sized businesses, and individuals, without requiring a proliferation of bespoke (but only marginally different) qualifications.

He also recommended that the UK Commission for Employment and Skills should work with employers to agree the future model for occupational standards, and that (In England) the same occupational standard should be used as the basis for apprenticeships, Tech Levels, and adult vocational qualifications.

The QCF

Whitehead followed Wolf in criticising QCF requirements, arguing that awarding organisations should be able to opt-out of certain of them, including requirements concerning unit format and unit sharing. He was also critical of the quality of QCF units based on existing NOS:

The use of the QCF has compounded the problem. NOS have to be rewritten into a standard QCF unit format and these units are added to the QCF unit databank. There is no quality assurance process to check these units reflect the initial NOS, are written clearly or are of an appropriate quality. The conversion of NOS into units adds a step to the development process. The approach of using a standard

QCF unit format was introduced so that individuals could change vocational qualifications and could transfer between awarding organisations more easily, avoiding unnecessary repetition of training. In practice, there is little evidence that the units system has resulted in individuals transferring between awarding organisations. Instead, the unit format has resulted in a databank of units not quality assured and used as building blocks for vocational qualifications. The format has also encouraged a “tick box” approach to curriculum and discouraged assessment that confirmed the overall standard had been reached.

(Whitehead, 2013, page 18)

He proposed that the weak link between NOS and QCF qualifications had led to proliferation, providing an example of a single NOS having been converted into QCF units that generated around 140 separate qualifications. His reforms would help to bring down the number of adult vocational qualifications on offer.

Curriculum leadership

Finally, Whitehead strongly supported concerns raised by the CAVTL report over the potential for negative washback on curriculum planning associated with highly prescribed standards and criteria. This level of prescription encouraged providers to treat discrete assessment requirements – detailed performance criteria – as though they laid out a coherent curriculum. Far less prescriptive standards would, he believed, encourage providers to think more carefully about curriculum design, particularly given the freedom it would offer to customise qualifications to the needs of local employers.

Whitehead insisted that the process for developing new standards had to be led by employers, making them more ambitious, aspirational, accessible, adaptable, and innovative. This brought the idea of ‘industrial partnership’ to the fore. Not only would employers be expected to lead the development of standards, they would be expected to influence qualification development too:

Awarding organisations should include employers from relevant sectors directly in the design and development of vocational qualifications, and training providers should bring in employers to support curriculum design and delivery.

(Whitehead, 2013, page 4)

Whitehead suggested that, over time, these employer-led partnerships would come together to take end-to-end responsibility for workforce development in their sectors.

VQ Reform Plan

Insights from all 4 of these reviews were integrated in the government's 'Reform Plan' for vocational qualifications (BIS, 2014). This report helped to clarify the complicated circumstances surrounding the reform of occupational standards, whereby:

- the UKCES continued to develop NOS on behalf of England and the devolved administrations
- 8 Trailblazer groups were developing new standards for apprenticeships in England
- the UKCES had begun to develop characteristics of higher-level occupational standards, along the lines set out in the Whitehead report

The Reform Plan report stated that:

In order to get maximum value from the effort that employers have put into developing new Apprenticeship standards, the Government believes that these should form the basis of any new National Occupational Standards that are developed. We are asking the UK Commission for Employment and Skills to make sure that any new NOS which are produced draw on the content of the relevant new Apprenticeship standard.

(BIS, 2014, page 12)

It explained that legislative change was required to facilitate the transition from apprenticeship frameworks to apprenticeship standards, and that frameworks would continue to be developed during the transition. The UKCES was charged with bringing the devolved administrations fully into the transition programme, to ensure that any new NOS would prove to be satisfactory across the whole of the UK.

In a subsequent statement of intent, the UKCES (2014) confirmed that it would establish greater clarity about what high-level outcome-based NOS might look like, anticipating that a varying spectrum of detail might be required with the move towards one NOS per occupation.

An article in 'FE Week' from December 2015 illustrated how the strained relationship between NOS and Trailblazer standards was becoming increasingly problematic (Lindford, 2015). Whitehead was quoted as warning that the Trailblazer process was "out of control" and there was a risk of NOS being bypassed entirely. He still believed that there was an important role for NOS as the (more detailed) foundation upon which the new apprenticeship standards and many vocational qualifications should be based. Conversely, a BIS spokesperson was quoted as saying that, although this was possible in theory, most Trailblazer groups had chosen a different approach.

Sainsbury report

The last report of significance to the future of the CASLO approach in England was the report of the Independent Panel on Technical Education, which was chaired by the former businessman and politician, Lord David Sainsbury (Sainsbury, 2016). The Panel had been established in November 2015 by Nick Boles, Minister of State for Skills, to advise on how to simplify and improve the quality of technical education in England.

Reflecting upon a century of failed reforms, which merely tinkered around the edges, Sainsbury explained that a central feature of an effective technical education system is “a well-understood national system of qualifications that works in the marketplace” (Sainsbury, 2016, page 6). He therefore made qualification reform the basis of his recommendations, insisting that this system should:

- be designed by government, but with “the knowledge and skills, and methods of assessment, for each qualification” (page 6) laid down by industry experts ¹⁵¹
- provide clear and simple routes into employment in specific occupations (that is, far fewer routes than currently available)
- be sufficiently flexible to allow learners to change routes, and to accommodate returning adults
- provide a preparatory transition year for students who are not yet ready to embark upon a technical education route post-16

The problem with the current system, Sainsbury argued, was that it was too complex and delivered the wrong skills:

Currently over 13,000 qualifications are available for 16-18 year olds, yet these often hold little value for either individuals or employers, although that may not be obvious until too late. At higher levels, too, technical education qualifications have too often become divorced from the occupations they should be preparing individuals for because there have been no, or only weak, requirements that they meet such needs.

(Sainsbury, 2016, page 8)

Calling for a fundamental shift in technical education, the report made 34 recommendations, which began:

¹⁵¹ Emphasising the importance of industry buy-in, he emphasised that the system would: “only work if industry takes ownership of the content and standards of technical education, and makes certain that companies adhere to them” (Sainsbury, 2016, page 7).

Recommendation 1: We recommend the Government develops a coherent technical education option which develops the technical knowledge and skills required to enter skilled employment, which leads from levels 2/3 to levels 4/5 and beyond, and which is highly valued because it works in the marketplace.

Recommendation 2: The technical education option should be recognised as having two modes of learning: employment-based (typically an apprenticeship) and college-based.

Recommendation 3: While it is necessary for government to design the overall national system of technical education, employer-designed standards must be put at its heart to ensure it works in the marketplace. A single, common framework of standards should cover both apprenticeships and college-based provision. These standards must be designed to deliver the knowledge, skills and behaviours required to perform successfully in specific occupations, not the narrower job role-focused needs of individual employers.

(Sainsbury, 2016, page 17)

Ultimately, these recommendations led to the development of college-based T Levels, which at Level 3 sat alongside academic A levels and employment-based Apprenticeships.

The Sainsbury report was clear that NOS should not be the basis for technical education qualifications in the new system, as these were: “derived through a functional analysis of job roles and this has often led to an atomistic view of education and a rather ‘tick-box’ approach to assessment” (Sainsbury, 2016, page 17). This underscored the principle that technical qualifications should be derived from the same (new) standards as apprenticeships. Learners following a college-based route would therefore develop essentially the same competencies as those working towards an employment-based apprenticeship.

Also of relevance to the CASLO approach, Sainsbury recommended that every technical education qualification should be assessed using realistic tasks and synoptic assessment – to test a student’s ability to integrate and apply their knowledge and skills – and recommended that all qualifications should include external assessment. Again, the idea of synoptic assessment was a response to concerns over the atomised approach to assessment associated with NOS-based NVQs, with its potential for negative washback impact on teaching and learning.

Recommendations from the Sainsbury report were accepted unequivocally (BIS & DfE, 2016). In the future, options for 16+ students would be either technical or academic. The future of existing qualifications that straddled both academic and technical pathways therefore hung in the balance.

Policy post-2010

Although none of these post-2010 policy reviews focused specifically on the CASLO approach, they embedded within TVET policy discussions concerns that had previously remained largely within the academic literature. Wolf focused primarily upon threats to qualification standards, which she associated with the mastery requirement, compounded by perverse incentives (linked to funding and performance tables) that risked undue lenience. Richard focused primarily upon the threat of negative backwash impacts upon teaching and learning, which included spending too much time assessing and not enough time teaching and learning, as well as the threat of not developing a sufficiently integrated, holistic competence. The CAVTL report also recognised these threats, although it did not necessarily accuse the CASLO approach, *per se*. It merely noted a tendency for teachers and trainers to treat CASLO qualification specifications as though they represented learning programmes, without appreciating that they were simply the foundation for planning curriculum and pedagogy. It concluded that colleges and training providers needed to reassert their ownership of curriculum and pedagogy. Whitehead echoed concerns regarding the risk of learners not developing a sufficiently integrated, holistic competence. So, too, did Sainsbury, a few years later.

For 3 of these reports – Richard, Whitehead, and Sainsbury – the risk of learners not developing a sufficiently integrated, holistic competence could be mitigated by basing apprenticeships and vocational qualifications upon (the same) new, short and easily understandable, employer-led standards. These new standards would capture the spirit of competence succinctly, rather than the letter of competence comprehensively. In 2017, responsibility for overseeing the development of these new standards fell to the Institute for Apprenticeships (IfA), which subsequently became the Institute for Apprenticeships and Technical Education (IfATE). As the IfA was launched, the UKCES was wound up. The NOS system continued to service the devolved administrations, albeit with no formal input from England.

Three of the reports – Wolf, Richard, and Sainsbury – recommended that assessment should be at least partly external. This was an indirect criticism of the CASLO approach, as it tends to be associated with continuous or phased centre-based or work-based assessment. The 3 reports argued that external assessment was necessary to ensure the consistent application of national standards, although the DfE subsequently characterised this more in terms of requiring a comparable level of challenge between vocational and academic qualifications, which was also part of the DfE justification for synoptic assessment. Whitehead also recognised the importance of externality, but argued that this could be ensured by effective external verification.

Both Wolf and Whitehead expressed dissatisfaction with the QCF on numerous fronts. Wolf argued that it was not servicing the needs of 14 to 19 vocational learners. Whitehead made the same argument in relation to adult vocational learners. The core characteristics of the CASLO approach were implicated in this. Having said that, it is worth noting that the reports were not critical of all CASLO qualifications. For instance, Wolf acknowledged repeatedly that Level 3 BTECs were valued in the labour market and by higher education, and provided very high positive returns to learners. The vast majority of BTECs were based entirely upon the CASLO approach at that point in time, as they had been for many years.

Post-2010 regulatory decisions

Ofqual began operating as a discrete entity in May 2008, albeit still technically located within the QCA. Once legislation came into force, in April 2010, Ofqual began operating as a fully independent regulator. Ofqual inherited the National Qualifications Framework and associated regulations from QCA, which governed GCSEs, A levels, NVQs, and other qualifications. It also inherited the Qualifications and Credit Framework, and published regulations for the QCF in August 2008. Ofqual began recognising awarding organisations and QCF qualifications, assuming that the vast majority of vocational qualifications that were regulated under the NQF would transfer to the QCF by the end of 2010.

The shift away from the NQF prompted a review of Ofqual's regulatory approach. Following a series of consultations, Ofqual adopted a new regulatory approach in May 2011. Its focus was on awarding organisations, and its intention was to ensure that all recognised awarding organisations became fully responsible for the quality, standards, and value for money of their qualifications. Two regulatory documents were central to this approach:

- the Criteria for Recognition of awarding organisations (Ofqual, 2011c)
- the General Conditions of Recognition (Ofqual, 2011d)

All regulated qualifications were now regulated under these conditions, although certain qualifications were also regulated under a series of 'Additional Regulatory Documents' that remained in force, including:

- the GCSE, GCE, Principal Learning and Project Code of Practice
- the NVQ Code of Practice
- the Regulatory Arrangements for the Qualifications and Credit Framework

This meant that NQF Statutory Regulations were superseded by the General Conditions of Recognition (GCR), while QCF regulations remained in force alongside

the GCR. The NVQ regulations remained in force as NVQs were gradually being transferred into the QCF.

Withdrawing the QCF

Although early evaluations had identified teething problems with the QCF (Ofqual, 2011b), and although the status of the QCF required careful consideration when developing Ofqual's new regulatory approach (Ofqual, 2009a), there was no suggestion prior to publication of the Wolf report that the fate of the QCF might hang in the balance. Even in 2013, the Minister for Skills continued to support the role of the QCF in supporting adults who required tailored learning programmes, as well as adults who required small, accessible, cost-effective units of learning, such as offenders and unemployed people (DfE & BIS, 2013).

Having reflected on the reports by Wolf, Richard, and Whitehead, and other reports too, Ofqual commissioned an internal review of the QCF toward the end of 2013. In July 2014, it released a consultation on withdrawing its regulatory arrangements, which included details of the internal review as an appendix (Ofqual, 2014b).

Influenced by the post-2010 policy reports, Ofqual's review recognised that the QCF was not achieving its intended positive outcomes and, worse still, had resulted in certain unanticipated negative consequences. This included concern that QCF regulations were incentivising the development and delivery of qualifications that were neither meeting the needs of the relevant sector nor were assessed appropriately. The review was also cognisant of the direction of travel of recent DfE policy decisions related to grading, synoptic assessment, and end-point assessment, none of which aligned to QCF design rules. The review also recognised Ofqual's new statutory objectives, which were not in force when the 2008 QCF regulations were being drafted, and which stressed that Ofqual's primary role was to uphold the validity of qualifications and assessments.

The most fundamental criticism of the QCF, which was highlighted by the review, concerned the relationship that it had established between units and qualifications:

For many, a qualification should add up to more than the sum of its parts in a way that a set of accumulated units does not. For a number of stakeholders, from the time when the QCF was launched, this approach was damaging and contributed to the destruction of established and well-regarded qualifications. Stakeholders who were involved in the development process talk about having to break down qualifications to try to 'shoe-horn' the components into the unit template in order to get the qualification onto the QCF. Many also take the view that there is something which is educationally flawed in this approach to the creation of qualifications and that in starting with the unit, what's lost is the sense of the whole qualification being worth more than the sum of its parts

(Ofqual, 2014b, page 57)

Embodying this anomaly, some of the organisations that submitted units to the QCF unit bank were not actually awarding organisations, and were not recognised by Ofqual. Indeed, Ofqual had no role in quality assuring units. Furthermore, around 10% of submitted units had not actually been used by an awarding organisation. Finally, although QCF rules of combination had been established to ensure the internal coherence of unit-based qualifications, they appeared not to be working, as too many qualifications appeared to be little more than a “bundle of units” with no relevance to employers and no value to learners (Ofqual, 2014b, page 68). The potential problems associated with introducing a heavily unitised system – which had been well rehearsed prior to the introduction of the QCF (see Unwin, 1999) – had now come home to roost.

In addition to other fairly fundamental concerns for the viability of the framework – including problems arising from unit sharing, limited evidence for the utilisation of credit transfer, and so on – the review identified more specific issues of direct relevance to the CASLO approach. These included:

- the risk that specifying units in terms of learning outcomes might have turned assessment into a mere box-ticking exercise
- the risk that specifying qualifications in terms of units (and, in turn, learning outcomes) might work against synoptic, end-point assessment
- the risk of over-assessment

The review was even firmer in its critique of the QCF mastery requirement:

For competence-based qualifications, and particularly those related to a licence to practise, the mastery model is not only common but many would consider essential. The often-quoted example is of the airline pilot; we all need to have confidence that she can land the plane as well as take-off and fly it. For other types of qualification, and there are many of them on the QCF, the mastery model is not appropriate and again raises issues about the rigidity and inflexibility of the QCF.

(Ofqual, 2014b, page 69)

The review also noted that the CASLO mastery requirement effectively proscribes certain assessment approaches that are normally premised upon sampling of qualification content, including multiple-choice tests.

Ultimately, the review concluded that the QCF regulatory arrangements were not fit for purpose, and should be withdrawn. Awarding organisations should no longer be constrained by highly prescriptive design rules. Instead, they should be required to develop coherent qualifications that would be judged primarily in terms of validity.

The consultation document indicated that Ofqual agreed with conclusions from the review and recommended that the QCF regulatory arrangements should be withdrawn. High quality QCF qualifications should continue to thrive – regulated solely under the General Conditions of Recognition – but invalid qualifications would need to be amended or withdrawn. Gone, too, would be the bank of units from which awarding organisations could draw. The new system would revolve around qualifications, not units, for which awarding organisations would be solely responsible:

From now on, we will be clearly placing validity at the centre of our approach to regulation: a qualification as a whole must be valid, not just the individual units within it.

(Ofqual, 2014b, page 9)

The consultation document also came down strongly against the idea of requiring all qualifications to adopt a mastery approach:

Our proposals on assessment will also make it possible to move away from the mastery approach required of all QCF-type qualifications and to provide for compensation. This will mean that for some qualifications, a student's real strength in one area may be able to compensate for comparative weakness in another. We judge that this is likely to have a beneficial effect on all students and for many types of qualification will result in fairer outcomes.

(Ofqual, 2014b, page 27)

In December 2014, Ofqual announced its post-consultation decision to withdraw the QCF regulatory arrangements (Ofqual, 2014c). Ofqual's proposals for implementing new arrangements included developing a new qualifications framework that would be less prescriptive and more descriptive (Ofqual, 2015a). This framework – which was to become known as the Regulated Qualifications Framework (RQF) – would encompass all regulated qualifications. Ofqual also announced its intention to withdraw all residual NVQ regulations. The RQF was introduced in September 2015 (Ofqual, 2015b). Regulated qualifications in England were no longer required to adopt any of the 3 core characteristics associated with the CASLO approach (related to outcomes, criteria, and the mastery principle).

Commitment to 'strengthen' VTQs

Evidence of cohort-level results improving steadily over time inevitably raises questions concerning grade inflation. Maybe the cohort is not improving after all? Maybe the qualification standard has fallen? During the noughties, steadily improving cohort-level results raised serious concerns over grade inflation at GCSE and A level. In 2012, Ofqual intervened to address these concerns (Newton, 2022). By

2015, similar concerns had begun to be expressed regarding steadily improving cohort-level results in Level 3 BTECs (HEFCE, 2015). Ofqual developed a complex statistical methodology for investigating this possibility, which triangulated data from a range of sources. The research explored outcomes for 4 cohorts of students between 2005 and 2015, which restricted the analysis to 'old-style' BTECs, that is, to BTECs that were based entirely on the CASLO approach.¹⁵² The report provided strong evidence of grade inflation, questioning the extent of genuine improvement in cohort achievement over time (Cuff, Zanini, & Black, 2018).

This research report was published in December 2018, not long after Ofqual had announced its intention to regulate VTQs (particularly those that featured in school and college performance tables) with "the same seriousness and focus as we do general qualifications" (Ofqual, 2018a, page 5). It speculated that the grade inflation might have occurred as a result of BTECs adopting the CASLO approach, with 100% centre assessment, relatively weak controls over qualification standards, and perverse incentives caused by accountability mechanisms:

These findings might be explained by differences in the marking/awarding processes that exist for 'older style' Level 3 BTECs and A levels. For example, while a compensatory approach is taken for A levels, 'older style' Level 3 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, page 14)

Indeed, the report speculated that similar issues might arise for other CASLO qualifications operating within similar contexts. Ofqual concluded that there was a general case for strengthening controls over internal assessment in VTQs, particularly under the pressure of school and college performance table accountability (Ofqual, 2018b).

This research was acknowledged by the DfE, which concluded that there was a likelihood that Level 2 qualifications were also vulnerable (DfE, 2019). The DfE welcomed the programme of work that Ofqual had set in train to strengthen controls in VTQs.

¹⁵² Many of these were subsequently reformed to comply with DfE performance table requirements.

With a view to harmonising its regulatory requirements with the DfE's pre-existing design requirements for performance table qualifications, Ofqual established a programme of work that focused specifically on Level 1/2 Technical Awards taught to 14 to 16-year olds in key stage 4. This led to a set of decisions that were to be operationalised through new Qualification Level Conditions (Ofqual, 2020). Of particular relevance to the CASLO approach, these new regulations specified that awarding organisations should use a numerical, mark-based approach to both exam and non-exam assessment components (albeit with an option to apply for exemption). Explaining its rationale for prohibiting use of the CASLO approach within centre-based assessment components, Ofqual noted:

In addition, we think that this will enable awarding organisations to have adequate control over marking judgements made in centres, not least as it will provide greater scope for any adjustments to the marking standard that an awarding organisation might seek to make through their moderation process.

(Ofqual, 2020, page 20)

This was not the first time that Ofqual had prohibited the CASLO approach for particular qualification types (by requiring numerical marking). For instance, in the T Level context, the draft 'Technical Qualification Conditions and Requirements' that accompanied Ofqual's consultation on rules and guidance (Ofqual, 2018c) specified that candidates' performances had, in each assessment, to be differentiated by the allocation of numerical marks. The same decision had been reached for Essential Digital Skills Qualifications (Ofqual, 2019). In theory, numerical marking increases the control that an awarding organisation has over qualification standards in 2 ways:

- consistently lenient or harsh centre-based marking can be tackled by moderating centre marks down or up accordingly
- where standards appear to be out-of-alignment from year to year, at the cohort level, this can be tackled by raising or lowering grade boundaries accordingly

Yet, it is also important to acknowledge that circumstances do not always permit awarding organisations to capitalise upon these controls, especially when relatively small cohort sizes compromise the use of statistical modelling.¹⁵³

¹⁵³ Of course, prescribing numerical marking also means prescribing a compensatory aggregation principle (as opposed to a mastery one). So, what might seem like a relatively minor technical assessment requirement is actually a more fundamental one, with implications for curriculum, pedagogy, and assessment, as well as for certificate interpretation and use.

Regulation post-2010

Ofqual's decisions to withdraw the QCF and to strengthen VTQs were made independently of government, although they clearly:

- echoed concerns expressed in the post-2010 policy reviews, and
- ensured that regulatory qualification requirements were appropriately aligned to existing departmental qualification requirements and policies

The decision to withdraw regulatory arrangements for the QCF (alongside residual NVQ regulations) meant that the regulator no longer required any regulated qualification to adopt the CASLO approach. If an awarding organisation was to adopt the CASLO approach in the future, then it would be their choice to do so (although this decision might be influenced by key a stakeholder, such as a professional body).

In a number of instances, however, Ofqual decided that awarding organisations should not have this choice, particularly where there might be a strong perverse incentive for lenience. This included Key Stage 4 Technical Awards, T Level Technical Qualifications, and Essential Digital Skills Qualifications. It is worth noting that this has not become a general policy stance following the withdrawal of the QCF. For instance, there is no requirement for numerical marking of centre-based assessments within recently published Qualification Level Conditions for Alternative Academic Qualifications (Ofqual, 2023).

One final point to note is how Ofqual followed Dearing and Wolf in not exploring the relationship between mastery certification and mastery learning, that is, in not discussing the potential value of mastery learning for students on general or vocational courses. The consultation accepted that for competence-based qualifications – particularly those leading to a licence to practice – mastery was often considered essential (Ofqual, 2014b). Similarly, where consultation responses addressed this issue, they made the same point, particularly for certain sectors like health and social care (Pye Tait Consulting, 2014). Yet, the idea of mastery as a philosophical principle in its own right was not discussed.

Down but not out

There are now fewer regulated CASLO qualifications (mid-2020s) than there were a decade ago. However, it is unclear exactly how many CASLO qualifications Ofqual currently regulates – whether full or hybrid – as this information is not collated centrally. Although the approach is prohibited for certain qualification types, it is still permitted in many instances, and the approach is still viewed positively by many awarding organisations, particularly those dealing with competence-based qualifications (akin to NVQs). Indeed, there are even isolated examples of outcome-

based approaches within current General Qualifications and national curriculum assessments, including the A level science endorsement of practical skills, and teacher assessment of writing at key stage 2. It is fair to say that the CASLO approach has fallen out of favour with policy makers over the past decade or so. But, while down, it is certainly not out.

Chapter 7. Conclusions

In this final chapter, we will consider lessons that we have learnt concerning:

- the CASLO approach
- its fitness for purpose, and
- TVET qualification reform more generally

The CASLO approach

At the outset of our research programme, we defined the CASLO approach in terms of qualifications with units that shared 3 core characteristics:

1. unit content is specified via learning outcomes
2. the unit standard is specified via assessment criteria for each learning outcome
3. to pass the unit, a learner must acquire all of the specified learning outcomes, which we refer to as the mastery requirement

This was a very pragmatic definition, based simply on our observation that very many regulated TVET qualifications shared these 3 core characteristics, even when they shared little else. We now know much more about the origins and evolution of the approach, and the following sections set out some of our key insights.

How to understand the approach

We begin by reflecting on how best to understand the approach, including:

- the significance of diversity within the CASLO family
- the historical contingency of the approach
- the challenge of characterising the approach
- the need to locate our understanding within a broader theory of qualifications

Diversity

In the report of our first investigation in this area, report 3, we attempted to characterise the ‘archetypal’ CASLO qualification (beyond its 3 core characteristics). We concluded that qualifications of this sort tended to be designed with flexibility in mind, to accommodate learners studying under different circumstances, within different locations, at different times, and so on. And we noted that their heavy assessment load tended to dispose them toward a phased, or continuous,

assessment model.¹⁵⁴ But it is fair to say that the very idea of an ‘archetypal’ CASLO qualification was tricky to articulate, given the wide variety of qualifications within this very broad family.

A more interesting question, perhaps, was why the 3 core characteristics became so prevalent across such a widely divergent landscape. One part of the answer to this question was quite straightforward: the features that comprise the CASLO approach were specified as accreditation criteria for both the NVQ framework and the QCF. However, another part of the answer was more subtle, suggesting that these frameworks were pushing at an open door. After all, as early as the 1970s, much of the sector had bought into the importance of outcomes and mastery when designing vocational and technical qualifications. As such, the 1980s NVQ model was continuous with the 1970s TEC and BEC models, just as the 1990s BTEC model was continuous with the 1980s NVQ model.

Significantly, although both 1980s NVQs and 1990s BTECs fully adopted the CASLO approach, it seems appropriate to locate these models at opposite ends of a continuum. At one end was the original NVQ model, which was neutral to teaching and learning approaches, on the assumption that a qualification ought to be achievable in a variety of different ways. At the other end was the original BTEC model, which was committed to a particular approach to teaching and learning, on the assumption that it was optimal for its targeted cohort.

The prevalence of distinct traditions of this sort makes it hard to learn lessons about the CASLO approach, per se, from historical analysis alone. For instance, where the approach came to be associated with progressive, student-centred teaching and learning – which was true for BTECs, GNVQs, and many other qualifications – it is hard to distinguish strengths and weaknesses associated with the CASLO approach from strengths and weaknesses attributable to broader aspects of their philosophy, design, development, and delivery.

The most important insight, though, is the simple fact that there is a lot of diversity within the CASLO qualification family. Different manifestations of the CASLO approach – including the NVQ model, the GNVQ model, the BTEC model, and the QCF model – are best understood as distinct nodes within a broad network of approaches circumscribed by the general CASLO definition. It is also important to appreciate that the CASLO approach itself is best understood as just one node within a far broader network of qualification models that emphasise the importance of outcome specification and outcome mastery.

¹⁵⁴ We also noted that many CASLO qualifications only awarded the passing grade, and that many were not heavily time constrained.

Historically contingent

Bearing in mind for how long the CASLO approach dominated the TVET qualification landscape in England, it would be easy for newcomers to fall into the trap of assuming that it must simply be the 'occupational way' of designing qualifications. Yet, even a cursory analysis reveals that this is not really true. Before the introduction of outcome-based models, college-based TVET qualifications were designed classically, and some still are today, both nationally and internationally. The approach makes a lot of sense in certain occupational contexts – where it certifies full occupational competence – but it is not the only game in town.

It is more appropriate to conclude that adopting the CASLO approach was historically contingent, in the sense of representing a particular response to a particular confluence of problems at a particular point in time. As just noted, factors such as qualification accreditation criteria were key to explaining why it became so dominant, and if these factors had not operated then history might have been quite different. England might, for instance, have explored the potential of a far wider variety of outcome-based and mastery-based approaches. We will return to this issue in the final report of our research programme (report 9).

Adoption of the approach was also historically contingent in the sense of riding the wave of a number of North American educational movements, related to objectives, mastery learning, and criterion-referencing. The concept of criterion-referencing became particularly influential in England during the 1980s. For a variety of reasons, it struck a chord with both scholars and politicians alike. Both general education and technical training qualifications changed in response to this criterion-referencing zeitgeist, but in different ways.

The exam boards were instructed to investigate making their exams more criterion-referenced. Having investigated this, they concluded that radical reform was not appropriate for GCSEs or A levels. Consequently, General Qualification reforms of the 1970s and 1980s tended to be more evolutionary than revolutionary. This included developing far more detailed syllabuses, which, in addition to content, included: clear statements of aims, assessment objectives, grade descriptions, specification grids for papers, advice for teachers and students, suggestions on teaching approaches, details of recommended reading, and so on. Consequently, syllabuses that had spanned perhaps 2 pages in 1960 spanned some 20 pages by 1980 (Kingdon & Stobart, 1988).¹⁵⁵

¹⁵⁵ Syllabuses came to be known as 'specifications' during this period, such that the idea of a 'specification' superseded the idea of a 'syllabus'. Having said that, because the idea of a specification has less everyday currency, and is far more generic, the term 'syllabus' is still quite

Although GQ reforms were extensive, TVET qualification reforms of the 1970s and 1980s were far more revolutionary. Recognising the inadequacy of planning curriculum, pedagogy, and assessment on the basis of meagre syllabus content lists, a new foundation for planning was devised. This involved specifying learning outcomes in detail, meaning that everyone involved in teaching, learning, and assessment could refer to the same, precise statement of the outcomes that students would be expected to acquire (rather than having to rely on a brief, ambiguous statement of what teachers would be expected to cover in their teaching). This distinction was captured in the idea that the new qualifications would be ‘outcome-based’ (learning-focused) rather than ‘input-based’ (teaching-focused).

Inputs ‘versus’ outcomes

In fact, although this distinction appears to tap into something important, consistent with the idea of a genuine revolution in TVET qualification design, it is conceptually problematic, and overstates the differences between outcome-based and classical approaches. One problem with the utility of the distinction is that it was invented to highlight a solution to a recognised design flaw with the classical approach: relying upon meagre syllabus content lists risked low-level examining, which risked low-level teaching and learning. The solution involved stating high-level (and low-level) intended learning outcomes explicitly, as the foundation for planning curriculum, pedagogy, and assessment. In short, the distinction was not invented to characterise 2 conceptually distinct approaches (outcomes versus inputs), it was invented to foreground the solution (outcomes) to a problem associated with the old approach. The ‘outcome’ idea was meaningful and useful, but not so much the ‘input’ idea.

A related problem with the utility of the distinction is that the new approach was intended to correct the old approach, not to oppose it. Note that the grids that Tyler introduced simply supplemented traditional content lists with information on what students needed to ‘do’ with that content. The outcome-based approaches adopted by both the TEC and the BEC worked in essentially the same way. The BTEC approach supplemented this with insights into effective teaching and learning, which included providing information and guidance on delivery approaches, indicative reading, links to related resources, and so on. In other words, outcome-based approaches were not fundamentally opposed to classical ones. And, despite how the distinction is sometimes portrayed in the literature, outcome-based approaches were not forced, as a matter of principle, to reject anything that might be construed as an ‘input’ to teaching – like syllabus content, learning programmes, and suchlike.

useful, and we sometimes use it even when discussing current arrangements. We distinguish both ‘specification’ and ‘syllabus’ from the idea of ‘curriculum’ planning, which goes beyond the requirements of a particular specification, integrating broader considerations related to the organisation of teaching, learning, and assessment.

The passage of time has rendered the ‘input versus outcome’ distinction even more problematic because, as we have just seen, the classical approach evolved too. The most significant development for classical qualifications in England was the inclusion of assessment objectives within all GCSE and A level specifications. They did not function in exactly the same way as learning outcomes in CASLO qualifications, but they clearly helped to elaborate the nature of the outcomes that students were expected to achieve while studying for a qualification.

We therefore conclude that the ‘input versus outcome’ distinction is neither very meaningful nor very useful nowadays. We recommend that it should no longer be used. This is not to reject the concept of an outcome-based qualification, which is specified principally in terms of learning outcomes, which are used directly for planning curriculum, pedagogy, and assessment. But there does need to be greater clarity and precision in articulating what an outcome-based approach might stand in contrast to (as the term ‘input-based’ fails to articulate this). This position behoves us to revisit our pragmatic distinction between the ‘classical approach’ and the ‘CASLO approach’ to qualification design.

If the CASLO approach is best understood as just one node within a broader network of outcome-based and mastery-based models – and if the classical approach is capable of incorporating features normally associated with outcome-based models – then is the ‘classical versus CASLO’ distinction still tenable, even if simply in relation to regulated qualifications in England? Well, it is certainly not the case that all regulated qualifications fall neatly into one of these 2 categories, depending on whether their designers adopted the CASLO approach or the classical one. Indeed, there are numerous hybrid qualifications nowadays, which incorporate both classical units (typically assessed externally, via mark-based exams) and CASLO ones too (typically assessed by centres, on an outcome-by-outcome basis).

Hybridisation became particularly significant during the mid-2010s, in the wake of Department for Education design requirements for any qualification that was to contribute to performance table calculations. These rules effectively proscribed adopting the CASLO approach for a certain proportion of each qualification. A number of qualifications that had previously been designed entirely according to the CASLO approach, including many BTECs, were reformed to incorporate classical units. On the one hand, hybridisation can be seen as a rational response to external pressures. On the other hand, it can be seen as an uncomfortable melding of quite different educational philosophies. The idea of hybridisation certainly begs questions that merit further research and analysis.

Although not all regulated qualifications in England can straightforwardly be categorised as embodying either the CASLO approach or the classical approach, we think that there are still enough qualifications that are far more like one of the 2 approaches for the distinction still to have meaning and utility.

A theory of qualifications

The classical approach was heavily influenced by a pragmatic tradition of examining for university matriculation, which became established in England during the second half of the 19th century. This tradition had a reputation for shrouding its methods, principles, and machinery in a cloak of mystery and secrecy (Wallis, 1927).

Well into the second half of the 20th century, General Certificate of Education exams continued to have an aura of mystery and secrecy about them – even in terms of their syllabuses – to the extent that Pearce felt a need to theorise the “information structure” of public exams, identifying 5 sources of insight potentially available to ordinary teachers (Pearce, 1972, page 28):

1. syllabuses (which were generally implicit, meaning that their interpretation relied heavily upon the second source)
2. professional expertise (that comes from being a subject specialist within the same community of practice as the syllabus writer)
3. past exam papers (which provided clues concerning how the chief examiner might sample from the syllabus)
4. student results (which indicated the adequacy of their preparation)
5. chief examiner reports, where available (which provided more or less detailed accounts concerning the strengths and weaknesses of student performances)

Pearce noted that syllabus implicitness presented a particular challenge to novice teachers, especially arts teachers, whose syllabuses were the least explicit.

But how explicit should a qualification syllabus (specification) be? This is surely a legitimate topic for a theory of qualifications to engage with. It is not a straightforward question to answer though. In England, this question became entangled within long-standing ideological debates related to control of the curriculum. From the outset, there was a tension between what universities wanted students to be able to demonstrate for matriculation purposes and what individual schools wanted to teach their students. This led to a proliferation of syllabuses from which schools were able to choose. Tension of this sort also discouraged exam boards from specifying syllabuses too explicitly, on the assumption that looser specification would enable teachers to follow different routes through any particular syllabus.

During the second half of the 20th century, government became increasingly involved in these debates between universities and schools. This pressure led to a reduction in the number of GCE syllabuses on offer, in an attempt to secure greater confidence in the comparability of standards (Tattersall, 2007). This also led to the development of A level common cores and GCSE national criteria – which specified the content that would need to be included within all syllabuses in a particular subject

area at each level – resulting in far tighter specification of syllabus content during the 1980s and 1990s. Their development also opened the door to greater central control over the content of exam board syllabuses.

In response to this pragmatic tradition – and in the midst of these ideological battles – the TEC, the BEC, and the NCVQ decided that their qualifications needed to be specified far more tightly than qualifications had previously been, drawing inspiration from outcome-based approaches that had been associated with North American curriculum theory. They encountered a certain amount of resistance to the idea of tighter specification, particularly from those who believed that government should not exert so much influence over the curriculum, and that control should ultimately reside locally. However, it seems fair to conclude that most of the resistance to outcome-based approaches in England, particularly within the academic literature, related to exactly how these far tighter specifications were articulated. Of particular concern was the risk that unitised specification of learning outcomes would detract from the idea of an integrated, overarching competence.

So, what does the theory that underpins qualification policy and practice in England have to say of relevance to this debate? Apparently, not a great deal. There seems to be a surprising lack of scholarship devoted to the aims and functions of qualification syllabuses (now known as specifications). In fact, there are good reasons to question whether there even exists a body of work that might legitimately be described as a coherent underpinning theory for this area of policy and practice. Again, we will return to this issue in our final report (report 9).

Why the approach was introduced

Reading some of the more extreme critiques of the first CASLO qualification of national prominence, the NVQ, we might struggle to understand why on earth such a radical reform had ever been contemplated in the first place. Yet, we should not underestimate the problems that outcome-based qualifications were introduced to help solve. It was not just that qualifications for off-the-job education and training were unsatisfactory, provision for on-the-job training was unsatisfactory too. Reform was certainly due, if not long overdue.

Serious problems

It is sometimes said that the NVQ model (and therefore the CASLO approach) evolved out of a need to certificate training schemes for young people who might otherwise have remained unemployed and become increasingly unemployable. The implication is that many of the problems associated with NVQs can be attributed to tailoring what was supposed to be an all-compassing qualification framework to the needs of a relatively small cohort of learners. There is certainly some truth to this

claim. Indeed, exactly the same criticism could be made in respect of the Qualifications and Credit Framework, which was also problematically tailored to the needs of a relatively small cohort of learners, in that instance, returning adult learners. As far as the TVET landscape is concerned, one size never fits all, so the key to system rationalisation must be to seek an optimal balance between generic and bespoke.

In fact, the features that comprised the NVQ model, including the CASLO approach, were not introduced purely to solve the problem of YTS certification. They were intended to address multiple, serious, long-standing, widely-recognised issues within the TVET landscape of the 1960s and 1970s. These included problems related to:

- on-the-job work-based training – that was not well specified, that was largely uncertificated, and that was highly variable in quality
- off-the-job college-based education and training – that was skewed toward book knowledge, that sometimes had limited relevance to employment, that was criticised for failing to target the high-level competencies that employers really needed, and that was plagued by wastage, retardation, and failure
- the apprenticeship system generally – that was rapidly falling out of favour with employers
- the many young people and adults who received no education or training beyond school or college

The outcome-based approach to qualification design proved to be particularly useful in relation to those aspects of occupational competence that were not directly associated with the kind of book knowledge that was often the focus in college. Within a well-established academic discipline, it is possible to imagine how a meagre syllabus content list might (just about) be capable of providing sufficient conceptual scaffolding to enable a subject specialist with considerable experience and expertise to create an appropriate programme of learning. Yet, this was not the situation that the Industrial Training Boards faced during the 1960s. The idea of writing specifications for on-the-job training along the same lines as a traditional book knowledge syllabus would have made no sense. So, they developed specifications defined in terms of outcomes, albeit often very narrowly defined ones, concerning the tasks that an employee would need to perform (task analysis).

The NCVQ believed that it could generalise this approach to embrace outcomes arising from both on-the-job training and off-the-job education and training. This would require a broader specification of the occupational role that an employee would need to perform. Learning outcomes were defined in terms of what competent performance of those roles looked like (functional analysis).

This approach left implicit that which a traditional book knowledge syllabus would traditionally have foregrounded, that is, the knowledge and understanding that

underpins competent performance. In retrospect, it seems fair to conclude that this was a serious mistake, as it appeared to signal that book knowledge was unimportant for completing an NVQ, and this mired the implementation process in controversy for years. NVQs might have been received more positively if they had adopted a less radical proficiency model, perhaps more akin to those that had been developed by the TEC and the BEC, which were also intended to help counter many of the serious problems identified above.

Although the CASLO approach does not necessitate centre-based assessment, it naturally gravitates towards it, given its emphasis on mastery learning and certification. Centre-based assessment became part of the solution to 2 of the major problems that affected existing qualifications, which typically relied heavily upon terminal external exams. First, it helped to address the problem of wastage, retardation, and failure, by breaking down the overarching proficiency model into more manageable elements of competence. The requirement that all learners would have to achieve each element of competence established more clearly the need for progression to be carefully monitored, to reveal any emerging gaps in learning for immediate intervention. Second, centre-based assessment helped to address the problem of exams not targeting the higher-level competencies that employers really needed. This involved specifying those outcomes more explicitly and encouraging teachers to adopt alternative assessment formats whenever necessary to assess them.

Variety of purposes

Reasons for adopting the CASLO approach were typically not made explicit independently of reasons for adopting broader features of each of the qualification models considered in preceding sections (NVQs, GNVQs, BTECs, and so on). Consequently, we made a particular effort to unpack the multiplicity of goals that qualification designers appeared to want to achieve by adopting the CASLO approach. We identified 3 distinct perspectives on qualification goals:

1. the certification perspective – to improve the technical quality of assessment (validity)
2. the educational perspective – to improve teaching, learning, uptake, completion, and so on
3. the sociopolitical perspective – to improve the structure of the TVET system

One of the more explicitly stated goals that led to the CASLO approach being incorporated within NVQs was to safeguard the technical quality of NVQ assessment (particularly via domain alignment). Yet, ironically, one of the most serious criticisms of the CASLO approach was that it threatened technical quality. For instance, it was argued that because written criteria cannot possibly express standards with sufficient

precision to make perfectly consistent judgements, assessors would be liable to judge students according to different standards, which evidence revealed to be the case (Wolf, 1995; 2011). If so, then was it always naïve to assume that the CASLO approach ought to have a positive impact on the technical quality of assessment? There are 2 points to bear in mind here.

First, validity requires an overarching integrative judgement concerning all of the empirical evidence and logical analysis that bears upon claims concerning the accuracy of interpretations arising from qualification results. Because there is no such thing as perfect assessment, we always end up making design trade-offs that enhance certain aspects of technical quality while accepting a hit on others. This is true whatever qualification model we might be considering.

Advocates of the CASLO approach were keen to ensure that assessors assessed all of the right things (to ensure construct representation) and none of the wrong things (to avoid construct-irrelevant variance). They tackled these 2 principal threats to validity by modelling the target proficiency at the heart of each qualification in terms of learning outcomes and assessment criteria. In doing so, they aimed to specify the qualification construct both comprehensively and authentically. Unfortunately, there is often a correlation between authenticity and judgemental inconsistency in assessment contexts. Pragmatically, a proponent of the CASLO approach would hope that they had engineered a situation in which the increment in validity arising from improved construct representation (in particular) would outweigh any decrement in validity arising from judgemental inconsistency. Empirical evidence would be necessary to substantiate this hypothesis, of course. But, the point is simply that it is possible, in theory, to construct a plausible validity argument in favour of CASLO approach, which trades off different sources of validity and invalidity.

Second, the argument in favour of adopting the CASLO approach needs to be understood in terms of attempting to improve qualification validity, and not in terms of attempting to achieve perfect validity. With its plethora of learning outcomes and assessment criteria, the approach provides a great deal of scaffolding for assessment judgements. The less experienced the assessor, the more important scaffolding of this sort becomes. NVQs, in particular, were intended to support on-the-job training, where assessment expertise could not necessarily be guaranteed. The need for scaffolding would presumably have been high in this context. In short, the CASLO approach was never intended to turn workplace assessors into infallible judges, but it was certainly intended to enhance the validity of qualifications that relied heavily on their judgements.

Although improving validity was one of the more explicitly stated goals that led designers to adopt the CASLO approach, many other goals, particularly educational ones, were far less clearly articulated. It is unclear why, especially as the goals that we associated with the educational perspective were obviously extremely important.

We identified 4 distinct educational goals:

1. domain alignment – to align curriculum, pedagogy, and assessment as closely as possible with the intended domain of learning (and therefore also with each other)
2. domain mastery – to ensure that all students achieve a satisfactory level of attainment across the full domain of learning
3. qualification efficiency – to make the process of becoming qualified as efficient as possible
4. domain personalisation – to enable the domain of learning to be tailored to the personal situation, interests, or needs of learners (or customised to meet the needs of local employers) ¹⁵⁶

Note that the domain alignment goal is simultaneously a certification perspective goal (to improve the quality of assessment) and an educational perspective goal (to improve the effectiveness of teaching and learning).

As will have become clear throughout this report, adopting the CASLO approach did not always result in goals of this sort being achieved. More work is required to understand both the circumstances that facilitate their achievement and the circumstances that frustrate them. More fundamentally, though, we need to consider whether these goals are still as attractive today as they might have been in previous decades. It seems likely that some will be, and it seems likely that some will not. Once again, this is an issue to which we will return in our final report (report 9).

How the approach was received

Because the CASLO approach was introduced via qualifications that were innovative and idiosyncratic in many different ways, it is hard to distinguish between observations that relate directly to the approach and observations that relate to broader aspects of their philosophy, design, and delivery. This is particularly relevant in relation to how the CASLO approach was received.

Perhaps the most important point to note is that the CASLO approach was largely imposed on the TVET qualification landscape by the extended machinery of government, via bodies like the TEC, the BEC, and the NCVQ (in association with the bodies that controlled qualification funding decisions). Furthermore, it was introduced as one component of a network of policies designed to rationalise the landscape, which involved challenging extant power structures and provision. The context was therefore necessarily one of conflict.

¹⁵⁶ We also distinguished minimal from radical domain personalisation, and we distinguished domain personalisation from what we labelled 'teaching and learning approach' personalisation and 'assessment format' personalisation.

While reactions to the new TEC and BEC awards were mixed, reactions to the NVQ model tended to be more extreme, and frequently more negative. The NCVQ was generally perceived to be imposing change rather than facilitating it. This imposition was received differently in different quarters.

To the extent that the awarding organisations had been the agents of qualification 'proliferation' in the past, they were part of the problem that the new NVQ framework was supposed to solve. Ultimately, they worked closely with the NCVQ to establish their new roles in the landscape, as did the relatively new BTEC validating body. Although relations were sometimes extremely strained – particularly between the NCVQ and the BTEC – this conflict should not be overstated in relation to the CASLO approach. The principal awarding and validating bodies had largely embraced the shift towards outcome-based qualifications significantly before the NVQ framework was rolled out. This prefigured a wider embrace of criterion-referencing during the 1980s, which attracted scholars and policy makers alike.

NVQs received a mixed reaction from colleges, training providers, and employers. Funding incentivised their uptake, and many stakeholders bought into the system on that basis. There were plenty of employers who did not buy into the system, however, raising questions concerning the extent to which it was really employer-led. Funding incentives also drove GNVQ uptake, although there was a genuine sense that GNVQs were serving an important new purpose for full-time students in schools and colleges, for whom neither general education nor technical training routes seemed appropriate. Serious implementation challenges affected how NVQs and GNVQs were received. The CASLO approach received a better reception as the high-level design template for other awards, including OCN and BTEC awards.

Many education scholars reacted negatively to the introduction of NVQs and GNVQs. Again, these were innovative and idiosyncratic qualifications, and the CASLO approach was not always the principal bone of contention. But, it was definitely part of the critique. Debate concerning the introduction of outcome-based and mastery-based approaches was sometimes so heated that it became characterised as a battle ground.

Protagonists included government-sponsored agencies, especially the NCVQ, plus a contingent of education scholars who saw value in the new models. Although NCVQ publications tended to stick to procedural details, the NCVQ pursued an active research agenda, and collaborated with supportive academics in the production of scholarly outputs. In addition to Gilbert Jessup's own book on 'NVQs and the Emerging Model of Education and Training' (Jessup, 1991), this also included collections such as:

- 'Competency Based Education and Training' (Burke, 1989)
- the 'Competence & Assessment' quarterly journal, including occasional compendia such as 'Competence & Assessment Compendium No.2' (ED, 1992)

Beyond the relatively small group of academics who defended the new outcome-based approach, which included John Burke at the University of Sussex, a slightly larger group of academics was constructively critical. This included scholars like Alison Wolf, Michael Eraut, and Phil Hodgkinson. However, a considerably larger contingent of scholars was heavily opposed, including Kenneth Marshall, Alan Smithers, and Terry Hyland. Although the NCVQ actively engaged with critical voices (Tim Oates, personal communication), these debates rarely made it into print. Indeed, it would inevitably have been hard for the NCVQ to commit sufficient time to engage with this rapidly expanding literature, while simultaneously rolling out the NVQ and GNVQ models and responding to implementation challenges.

The hostility of the radical NVQ-GNVQ critique has been commented upon in the literature (see, for example, Bates, 1995; Ecclestone, 1997; Hargraves, 2000). No doubt passions were raised by factors beyond the underlying conceptual issues at stake, for instance, heavy encroachment by government into matters that would previously have been negotiated between awarding organisations, education providers, and professional bodies.¹⁵⁷ The fact that the NCVQ was an agency of the Employment Department, with limited involvement from the Education Department, might well have aggravated concerns. But it is also important to remember that outcome-based approaches were introduced in England at a time when debates over control of the curriculum would have been at their height. In this context, the very idea of prespecifying educational outcomes would have raised the hackles of anyone in favour of teachers retaining professional autonomy over such matters.

In this context, we noted that the NVQ literature includes strands that are not simply extreme, but problematically so. This includes the claim that the NVQ model – and outcome-based approaches more generally – are inherently behaviourist and therefore fundamentally flawed and unworkable. This claim has been repeated so many times in the literature that it has effectively become a matter of TVET dogma. We need to take this seriously, of course. If it is literally true that outcome-based approaches are fundamentally flawed – as a consequence of fundamental flaws within behaviourism itself – then this behoves the regulator to ban them from the TVET qualification landscape. Period.

Having considered the most radical strands of this academic critique in some detail, we did not reach this conclusion. For instance, on the particular issue of the supposed behaviourist roots of outcome-based approaches, we concluded that certain of the most radical claims were simply without foundation, including the idea that the NCVQ based the NVQ model on behaviourist learning theory. Other criticisms we judged to be unduly extreme, based upon a misleading account of the

¹⁵⁷ Concerns such as these can be seen in the literature that attempts to unpack the sociopolitical goals underlying the introduction of the CASLO approach (for example, Young, 2008).

development of outcome-based approaches in North America, and to some extent on a misleading account of the nature of the NVQ model. We exit this debate with lingering concerns over the legacy of the radical critique.

Its trajectory

Part of the empirical argument against the claim that outcome-based approaches are fundamentally flawed – and therefore inherently unworkable – relates to the trajectory of the CASLO approach in England. It is certainly fair to say that the CASLO approach was imposed on the TVET qualification landscape by the extended machinery of government, for instance, via accreditation criteria for the NVQ and QCF frameworks. As such, the CASLO approach did not simply ‘evolve’ in a survival of the fittest sense. Indeed, its survival was often secured by funding levers, which incentivised uptake of CASLO qualifications.

However, it is not fair to say that the origins and evolution of the approach can straightforwardly be explained in terms of central enforcement. The NCVQ had limited statutory powers, and the levers available to the government, including funding levers, could have been pulled more forcefully than they actually were. The truth no doubt lies somewhere between the extremes of central enforcement and natural selection.

Although the NVQ model remained controversial until its official demise, it never abandoned the CASLO approach. Furthermore, the fact that NVQs survived for a long time suggests that the model must have got something right (or, at least, not entirely wrong) and, despite sustained criticism, many employers remained reasonably content with its assessment approach (West, 2004).¹⁵⁸ While the post-2010 official policy reviews, including the Wolf report, raised important questions related to qualification standards, it is also the case that NVQs enabled many learners who might never have completed a qualification to obtain one. This stood in stark contrast to many TVET qualification suites that preceded NVQs, which had been plagued by drop out and failure. NVQs also provided a viable solution to the problem of certificating off-the-job training, which had not been adequately addressed prior to their introduction.

The GNVQ story is more complicated as far as the CASLO approach is concerned. Like NVQs, GNVQs were plagued by implementation challenges, such as assessment overload. Unlike NVQs, however, GNVQs were gradually redesigned to move away from the CASLO approach. The final iteration of the GNVQ model was the AVCE, which could no longer be described as a CASLO qualification. AVCEs were criticised as being neither seriously vocational, nor consistently advanced, yet

¹⁵⁸ Note also that, while England moved away from the NVQ model, Scotland has retained its SVQs.

their assessment regime still remained excessively complex, bureaucratic, and hard to understand. It was not until almost all of the remnants of the CASLO approach had been eliminated from the model that the qualification ultimately failed.

As GNVQs failed, so BTEC Nationals began to thrive. The BTEC family inherited new cohorts of learners from GNVQ schools that would not previously have considered a BTEC route. The BTEC model that thrived during the 2000s embodied the CASLO approach in full, in contrast to earlier models which had adopted somewhat different outcome-based approaches.

Into the 2010s, the CASLO approach was seen as key to the success of the QCF, so it was established as a formal design requirement for all accredited units and qualifications. By the mid-2010s, it had become clear that the CASLO approach was now the high-level design template for the vast majority of regulated TVET qualifications in England.

Central control

The origins and evolution of the CASLO approach are inseparable from the story of increasing central control over the TVET landscape in England. Back in the 1950s and 1960s, government actively influenced these systems, for example, through membership of the Joint Committees that oversaw the original National and Higher National awards. Yet, it was primarily the awarding, examining, and accrediting bodies that oversaw and co-ordinated these systems. However, precisely because there were so many bodies with responsibilities of this sort, there was no single national system, and the systems that coexisted were not co-ordinated. From the 1970s, government increasingly assumed control of TVET qualification systems in England, in an attempt to rationalise and simplify provision. It operated through a succession of independent non-departmental public bodies, which included the:

- Technology Education Council (1973 to 1983)
- Business Education Council (1974 to 1983)
- Business and Technician (Technology) Education Council (1983 to 1993)
- National Council for Vocational Qualifications (1986 to 1997)
- Qualifications and Curriculum Authority (1997 to 2008)
- Ofqual (2008 to present)

These organisations operated alongside a succession of bodies with related responsibilities for co-ordinating training and training standards at a national level, as well as alongside the existing awarding, examining, and accrediting bodies. Government departments also exerted control via qualification funding rules and, more recently, via school and college performance table rules.

These non-departmental public bodies influenced the structure of qualification systems in England, particularly through the design of qualification frameworks. However, they also influenced the design of the qualifications that populated those frameworks, through qualification (and unit) accreditation criteria. Both NVQ and QCF accreditation criteria established the CASLO approach as the national approach to designing TVET qualifications.

As the approach was rolled out nationally, through NVQs and GNVQs in particular, many problems were encountered. During the mid-1990s, a number of policy reviews were commissioned in response (including the Beaumont review, the Capey review, and the Dearing review). While acknowledging the scale of these implementation challenges, these reviews largely supported continued adoption of the CASLO approach, despite high-profile criticism of the model itself. A minor exception to this was the Dearing review, which argued that the concept of mastery was appropriate for NVQs but not for GNVQs (the Capey review expressed similar, but more nuanced, reservations concerning the mastery requirement for GNVQs, despite strongly supporting continued adoption of an outcome-based approach).

The CASLO approach continued as the national approach to designing qualifications for the technical training route, located at the heart of the NVQ model. This contrasted with the national approach to designing qualifications for the general education route, which continued to be the classical approach. Interestingly, though, the national approach to designing qualifications for the applied education (middle) route did change. The final GNVQ model (the AVCE) largely rejected the CASLO approach. And the national middle route qualifications that followed in its wake also adopted a classical approach (Applied A levels, the Diploma qualification, T Level Technical Qualifications). At the same time, however, the BTEC, which was no longer a non-departmental public body – meaning that BTECs were no longer a ‘national’ qualification as such – fully embraced the CASLO approach. So, the middle route, from the 2000s onwards, included qualifications based on both approaches.

Somewhat ironically, soon after the system had formally embraced the CASLO approach as the high-level design template for all QCF qualifications, it became a matter of concern within a number of official policy reviews (from the Wolf review to the Sainsbury review). None of these post-2010 reviews focused primarily on the approach, but they all raised CASLO-related concerns that were judged to be sufficiently important to be acted upon by government and its agencies in various ways.

The Department for Education responded to concerns raised in the Wolf report by specifying design rules for any qualification that was to be counted within a performance table calculation (from 2016 onwards). These included requirements for a certain proportion of external assessment and for the inclusion of synoptic assessment. This effectively ruled out the CASLO approach for a certain proportion

of each performance table qualification, and led to some qualifications adopting a hybrid approach, with both CASLO units and classical ones too.

The approach to certificating apprenticeships changed radically in response to the Richard review. This included rejecting extended, or continuous, assessment in favour of a terminal independent End-Point Assessment model. This represented a significant shift away from the CASLO approach – as it had been operationalised within the NVQ model – albeit without rejecting the articulation of learning outcomes or the idea of mastery.

In the wake of concerns expressed by the Wolf report and the Whitehead report, Ofqual reconsidered its approach to regulating TVET qualifications under the QCF. In 2015, it withdrew these regulations, subsequently regulating the vast majority of VTQs solely through its General Conditions of Recognition. As these conditions applied to all regulated qualifications, including GCSEs and A levels, they made no reference to the CASLO approach. From 2015, Ofqual no longer required any regulated qualification to adopt the CASLO approach (which remains true today).

Given this account, it is fair to say that the CASLO approach has fallen out of favour with policy makers in recent years. It is officially proscribed as a design template for certain qualification types. And for many other qualifications it is no longer required. Having said that, there are still many regulated CASLO qualifications in England, which is why we described the state of the CASLO approach as ‘down but not out’.

Fitness for purpose

We now turn to the issue of what lessons we might be able to learn from preceding chapters concerning the fitness for purpose of CASLO qualifications and outcome-based qualifications more generally. Unfortunately, it is not straightforward to reach conclusions of this sort. First, and most obviously, this project was not designed as an evaluation of the CASLO approach, in relation to the goals that tend to drive it. Indeed, part of the project rationale was to develop clearer insights into the nature of these goals, to provide a more solid foundation for subsequent evaluative work.¹⁵⁹

Second, as we have just seen, it is hard to distinguish strengths and weaknesses associated with the CASLO approach from strengths and weaknesses attributable to broader aspects of the philosophy, design, development, and delivery of the qualifications that we have studied. NVQs, GNVQs, and BTECs all incorporated

¹⁵⁹ Nor does it answer the even more complicated comparative question of the pros and cons of the CASLO approach versus the classical approach in relation to the different goals that they set out to achieve. The fact that CASLO qualifications are often designed to achieve quite different goals – mastery learning being an important case in point – would certainly complicate a comparative analysis of this sort.

innovative and idiosyncratic features in addition to the 3 core CASLO characteristics (outcomes, criteria, and mastery). Furthermore, the implementation failures that bedevilled NVQs and GNVQs make it hard to judge the viability of their underlying models on the basis of evidence from rollout alone.

Caveats aside, the preceding sections have clearly illustrated a variety of threats and tensions, and raised numerous questions related to the CASLO approach. After highlighting issues of this sort, we will consider what we might be able to conclude from this strand of our research programme concerning fitness for purpose.

Note that the following sections make repeated reference to issues raised by Alison Wolf. Wolf is one of the most significant figures in the field, having produced one of the most important scholarly texts (Wolf, 1995) and one of the most influential policy reviews (Wolf, 2011). Her reflections on fitness for purpose in relation to NVQs, GNVQs, and QCF qualifications – which have strongly influenced policy making in recent years – therefore provide an important point of reference. They are complemented by reflections from other scholars and policy reviewers.

Threats

We can classify threats that have been linked to the CASLO approach into one of 3 broad categories: threats to viability, threats to the quality of assessment, and threats to the quality of teaching and learning. Under each of these headings, the following subsections capture a range of issues that affected NVQs, GNVQs, BTECs, and other CASLO qualifications, and that help to explain why the approach began to fall out of favour with policy makers during the 2010s.

Threats to viability

Perhaps the most predictable threat associated with the CASLO approach derives from its requirement for exhaustive (all-encompassing) assessment. Because CASLO qualifications are intended to certify domain mastery, they require evidence that each and every specified learning outcome has been achieved to required standards. This tends to make the assessment process burdensome. Summarising experiences during the early 1990s, as NVQs and GNVQs were being rolled out, Wolf observed that every observer of the system – whether official body or independent researcher – agreed on the “sheer quantity of assessment” that teachers faced (Wolf, 1998, page 433).

For instance, evidence arising from her own 1994 investigation into GNVQ grading concluded that the average amount of time spent on assessment-related activities by each GNVQ teacher was 13 hours per week, with wide differences between teachers largely explicable in terms of class size. To the extent that the CASLO approach blurs distinctions between assessment, teaching, and learning – with its heavy

emphasis on formative assessment – this would not necessarily have been time spent unproductively. But the administrative load associated with portfolio assessment on this scale is undoubtedly challenging for both students and their teachers, trainers, and assessors. Indeed, Wolf argued that an important factor in explaining high GNVQ non-completion rates was the failure of many students to sustain a steady rate of portfolio completion.

Threats to the quality of assessment

Although the CASLO approach was often introduced to help secure the quality of assessment – as part of the rationale set out in the domain alignment goal – evidence from implementing CASLO qualifications illustrates significant threats too.

Inconsistent judgements

Wolf expressed even more concern over the threat of inconsistent judgements than over the threat of assessment burden. This line of criticism was developed across much of her research and analysis into NVQs and GNVQs during the 1990s, and it was also emphasised in her 2011 policy review. She claimed that evidence of inconsistent judgements challenged what she considered to be a widespread false belief concerning criterion-referencing:

Correspondingly, it is believed that it can provide – indeed, that its use guarantees – information about a candidate’s competence (skills, knowledge, etc.) that is *substantive and specific, and highly reliable*.

(Wolf, 1995, page 54)

It is debatable the extent to which protagonists of the approach genuinely believed that criterion-referencing guaranteed highly reliable judgements.¹⁶⁰ Yet, it is not debatable that NVQ and GNVQ judgements often exhibited limited reliability (see, for example, Eraut, Steadman, Trill, & Parkes, 1996).¹⁶¹ Furthermore, it is hard to argue with Wolf’s more general conclusion that written statements alone are insufficient for ensuring judgemental consistency. Wolf proposed that effective assessor networks – active communities of practice – were key to ensuring the consistent application of assessment criteria (Wolf, 1995).

¹⁶⁰ Wolf suggested that this was implicit in the relative lack of attention paid to this threat by those responsible for NVQ and GNVQ policy and practice.

¹⁶¹ A QCA evaluation of BTEC awards also raised concerns over inconsistent assessor judgements (QCA, 2005), as did a couple of Ofqual investigations (Ofqual, 2010a; Cuff, et al, 2018).

Deficient judgements

A different kind of threat to the quality of assessment was emphasised in the Richard review. This related to the risk that certification failed to capture information related to the integration and organisation of the elements that comprise occupational competence. By design, the CASLO approach deconstructs the overarching competence construct into discrete learning outcomes located within discrete units. If there is significantly more to the 'whole' (occupational competence) than might be captured by assessing individual 'parts' (discrete learning outcomes) then this presents a threat to assessment quality. This concern was echoed in Ofqual's 2014 evaluation of the QCF.

Threats to the quality of teaching and learning

Although less prominent during the 1990s, concerns over the potential for backwash impacts on teaching and learning became more prominent during the 2000s, and featured within the post-2010 TVET policy reviews. The need to shift attention from assessment back to curriculum and pedagogy was emphasised by the 2013 CAVTL report.

Superficial learning

The threat of superficial learning is closely related to the threat of deficient assessment judgements, stemming from the same phenomenon of discretely specified learning outcomes. The policy reviews written by Richard, Whitehead, and Sainsbury all expressed concern that the occupational competence acquired while studying for a CASLO qualification might end up being insufficiently integrated or holistic. The decision to specify learning outcomes far more generally, and succinctly, within new apprenticeship standards was a response to this threat of negative backwash impact from detailed CASLO specifications, as was the requirement for synoptic assessment within performance table qualifications.

Insufficient learning

If CASLO qualification implementation results in undue assessment burden, then this can lead to one of 2 possible consequences. The first is that additional time may need to be committed beyond the period already allocated to teaching and learning. The second is that the additional time required for assessment may eat into the allocated period. If so, then there will be less time available for teaching and learning. Undue assessment burden has often been linked to the CASLO approach, although often without unpacking its consequences in any detail. However, certain policy reviewers, including Doug Richard, expressed specific concerns over the threat of reduced teaching and learning time.

Tensions

Threats to assessment quality led to calls for greater rigour in relation to NVQ and GNVQ rollout. Reflecting on conclusions from the Beaumont and Capey reviews, Dearing noted that their: “greatest concerns related to the whole business of assessment, and at the heart of this lies the need to establish assessment that is rigorous” (Dearing, 1996, page 75).

Unfortunately, particularly in relation to NVQ implementation, this resulted in a fundamental tension between the need for qualification rigour and the need to be responsive to the demands of those who were charged with implementing the system, especially when (some but not all) employers valued simplicity and accessibility above rigour. This tension played out in numerous ways over the decades.

For instance, Beaumont was critical of the process of developing standards for NVQs, which he believed had been “marred by complex, jargon ridden language” (Beaumont, 1996, page 13). He argued that standards needed to “be written for employers” and that they “must all be in plain English” (both quotations from Beaumont, 1996, page 5). Yet, while system designers recognised these concerns, some were wary of what they saw as dumbing down the approach:

[The standards] sometimes include jargon – technical shorthand which is understandable only to occupational specialists. This is essential if the standard is to be used properly by the people who design and implement the systems based on the standards. If the standard is informal and colloquial it may be vague and often useless.

(Mansfield & Mitchell, 1996, page 154)

Tension of this sort is hard to reconcile. Assessment is a technical process, and the argument for expressing its requirements technically is persuasive. Yet, if this technical nuance proves to be inaccessible to those with primary responsibility for conducting the assessment, then it is not going to ensure rigour anyhow.

The tension between rigour and responsiveness played out even more starkly in relation to functional analysis itself, which had been introduced and advocated by the NCVQ and the Training Agency during the late-1980s. A decade or so later, when the QCA took over from the NCVQ, it sought views on how effectively the NVQ system was operating, particularly from the perspective of employers. The QCA was keen to make NVQ structures more flexible, to reduce bureaucracy, and to improve ease of use. Under new regulations that were published in 2000, functional analysis – and the detailed prescription that it entailed – was no longer formally required.

A decade or so later, the system for developing National Occupational Standards was reviewed by the UKCES with a view to improving its quality. It put functional

analysis firmly back on the table as a formal requirement of its new quality criteria, although its approach was not quite as stringent as the original methodology. Subsequently, the Sainsbury review rejected the very idea of detailed prescription that lay at the heart of functional analysis, and employers were invited to play a more significant role in developing standards.

Questions

After more than 3 decades of implementing the CASLO approach, there still remain numerous technical issues that would benefit from further research and development.

Learning outcomes

Plenty of questions remain to be answered concerning how best to specify learning outcomes for CASLO qualifications. These include how to tailor learning outcomes to the needs of particular qualifications, as well as more general questions related to learning outcome format, content, and grain size.

Format

The most important lesson to learn from NVQ implementation was that it matters greatly how learning outcomes are specified. The NVQ proficiency model was unusual in being specified purely in terms of the achievements that comprise performing an occupational role competently. Although this model made sense, in theory, it lacked credibility for many users. By leaving implicit exactly that which traditional qualifications had foregrounded – namely underpinning knowledge and understanding – it left itself open to widespread misunderstanding. It was caricatured as (and ridiculed for) shunning underpinning knowledge and understanding.

Although there are general lessons to learn concerning the consequences of adopting one or another format for specifying learning outcomes, questions still remain concerning how best to tailor learning outcome formats to the needs of particular qualifications, related to their purposes, cohorts, and contexts. Perhaps the most fundamental outstanding question concerns the relationship between outcomes that are more knowledge-like and outcomes that are more performance-like. Should they be specified in similar formats or quite differently? Should they be specified together or separately? The idea of certificating an integrated overarching competence seems to argue in favour of using essentially the same format within a single qualification. Yet, this was exactly the principle that underpinned the NVQ model, which was heavily criticised. In response to this critique, knowledge-like outcomes ended up being assessed quite independently, within Technical Certificates, which was also not ideal. The issue remains to be resolved.

Content

NVQs were also criticised in terms of the content of their learning outcomes. A major bone of contention concerned whether NVQ outcomes ought to be limited to, or be far broader than, specific occupational roles. By limiting NVQ standards to the parameters of specific occupational roles, the NCVQ was criticised for discouraging broader learning, and this criticism was not fully addressed by adding core skills requirements. It is important to remember, however, that concerns of this sort are not directly related to the CASLO approach, as it would be entirely possible to specify a far broader education and training programme in terms of learning outcomes and assessment criteria.

Related to the issue of outcome format, NVQs were accused of shunning traditional syllabus content, which seemed to be implied by their lack of explicit reference to underpinning knowledge and understanding. As we have discussed in some detail, this was technically untrue of NVQs, and it is more obviously false as a criticism of many other CASLO qualifications.

Indeed, the basic idea of an outcome-based approach – as epitomised in the work of Ralph Tyler – is that it is intended to provide a more detailed specification than would be provided by syllabus content alone, to provide a more comprehensive and authentic foundation for planning curriculum, pedagogy, and assessment. We saw how TEC and BEC awards manifested exactly this idea, transforming traditional syllabus content lists into explicitly stated learning outcomes. We also saw how later BTEC awards adopted a variety of different approaches to specifying both syllabus content and learning outcomes. In other words, the idea that outcome-based qualifications are somehow inherently opposed to the idea of syllabus content is straightforwardly wrong.

However, practical and theoretical questions do remain concerning the relationship between syllabus content and learning outcomes. For instance, are there qualification purposes, cohorts, or contexts for which it is important to dial down the central specification of syllabus content, and to rely instead upon more generic and therefore more widely applicable learning outcomes? This relates to the goal of (minimal) cross-context domain personalisation, which has been linked to the CASLO approach over the years. So, when is this flexibility a good thing, and just as importantly, when might it become problematic?

Finally, to the extent that central specification of syllabus content is deemed to be important for a particular CASLO qualification – as has often been the case – questions remain concerning how best to specify it. For instance, is it best to overlay outcomes directly on content (as per Tyler) thereby explaining what learners are expected to ‘do’ with the content? Or is it best to specify content quite separately?

Grain size

The idea of grain size is intended to capture the amount of detail in which learning outcomes are specified. This highlights another critical tension: the greater the detail in which learning outcomes are specified, the greater the potential for clarity, but also the greater the risk of mis-specification, where the proficiency in question can no longer be represented meaningfully or usefully in the level of detail being attempted.

Identifying the right grain size is one of the most important of all design decisions, if not the most important. Remember that the basic idea of an outcome-based qualification is to provide more detail than a classical qualification. This involves explicating what we want learners to be able to 'do' with the syllabus content to which they are exposed. Or, in the words of the authors of the most recent articulation of Bloom's Taxonomy, this involves making "general and abstract learning goals more specific and concrete" (Anderson & Krathwohl, et al, 2001, page 12). The more clarity we can provide, the better. But the pursuit of clarity forces us to consider when enough is enough. That is, we need to identify the point at which attempting to provide further detail becomes counterproductive, and clarity is lost.

This has always been the trickiest question for advocates of criterion-referencing to answer. Experiences from both England and overseas teach us that this typically involves trial and error. Reflecting on both the international literature and experiences in England, Wolf referred to the threat of "a never-ending spiral of specification" (Wolf, 1995, page 55). From a purely practical perspective, overstepping the mark is counter-productive because it results in specifications that are too burdensome to be useful, and that consequently fall into disuse.

In England, pitching learning outcomes at the right level of detail proved to be challenging right from the outset. The BEC was able to learn from the experience of TEC awards, and pitched its original outcomes at a slightly higher level of generality. Subsequent generations of BTEC awards experimented with differing levels of detail. The NCVQ methodology for specifying standards was supposed to culminate in outcomes that were fairly broad. Yet, in practice, these outcomes were often specified quite narrowly. The addition of range statements also led NVQs down the route of increasingly detailed specification.

Wolf has argued that the goal of "total clarity" means that outcome-based approaches are inherently doomed to overstep the mark, as written specifications can never deliver total clarity no matter how detailed they become (Wolf, 1995, page 54). Yet, as noted earlier, this seems to be an uncharitable reading of these approaches, which are fundamentally pragmatic, and which therefore aim for optimal clarity, not total clarity. The Goldilocks principle seems to be apposite here.

The question of how to specify outcomes at the right level of granularity, or detail, does not have a straightforward answer. Again, it is likely to depend to some extent

on the purposes, cohorts, and contexts for which a particular qualification is to be designed. The international literature provides useful insights (for example, Bloom, et al, 1971; Popham, 1994; Mager, 1997; Gronlund & Brookhart, 2008; Cedefop, 2017).

Mastery

We turn now to the question of whether there might be a threshold beyond which it becomes inappropriate to attempt to specify the proficiency model in more detail, for fear of mis-specifying it. Under the CASLO approach, this links directly to the idea of mastery, because the smaller the grain size, the greater the number of parts into which the overarching competence is deconstructed, and therefore the larger the number of elements that need to be mastered independently. This also relates to nature and function of assessment criteria, which also add to the mastery burden.

The BTEC grappled with the grain size challenge when transitioning from the BEC model (Generation 1) to the BTEC model (Generation 2). Instead of nesting lower-level outcomes within higher-level ones, the G2 model retained only the higher-level outcomes for which it also provided a list of indicative content. So, rather than prescribing the exact set of knowledge, skills, and understanding that comprised each higher-level outcome, this content list provided a broad indication of the sort of knowledge, skills, and understanding that would be relevant to achieving it. This meant that the higher-level outcome could be achieved in different ways, facilitating (minimal) cross-context domain personalisation.

This illustrates the more general point that, unless a learning outcome is specified in extreme detail, there are likely to be different ways of achieving it, and sometimes very different ways. It seems reasonable to suggest that, the less detail in which a learning outcome is specified, the greater the number of ways in which it might legitimately be said to have been achieved. One of the fundamental goals of the CASLO approach is to reduce ambiguity by providing greater detail concerning the elements that comprise the targeted proficiency. Yet, the greater the detail provided, the greater the risk that the model mis-specifies it, by attempting to deconstruct elements that cannot meaningfully be split into further parts, or by enforcing the mastery requirement at a level where the amount of detail is too great for the concept of mastery to be meaningful.

Related to this point, Wolf discussed the “inherent variability of the contexts in which competence is tested and displayed” such that a learner might genuinely be able to demonstrate competence in a number of contexts yet not in others (Wolf, 1995, page 68). This suggests that competence is not actually a binary concept, and that even

the concept of mastery must have an element of compensation, or perhaps charity, built into it.¹⁶²

Perhaps, as a general rule-of-thumb, the idea of mastery makes less sense and becomes less useful the smaller the grain size with which outcomes are specified? If so, then any CASLO qualification designer will need to be able to judge the borderline beyond which the benefits of disambiguation (by specifying an outcome in greater detail) are outweighed by the risks of mis-specification (by implying that each and every additional detail is indisputably critical to the outcome). Beyond this borderline, it would no longer be profitable to characterise a higher-level outcome as the sum of its lower-level parts.

The discussion of inherent variability prompts a related question concerning the function of assessment criteria within CASLO qualifications. Should they function definitively (such that satisfying all of the criteria defines having mastered the learning outcome) or should they function indicatively (providing an indication, but not a prescription, of what it means to have mastered the learning outcome)?

QCDA guidance on the QCF insisted that assessment criteria had to function definitively – that is, each and every criterion that was specified for a learning outcome had to be satisfied – which was also how NVQ performance criteria functioned. Perhaps, though, for certain qualifications, it might be more meaningful and useful for assessment criteria to function indicatively? The Eraut report appears to have reached the same conclusion in arguing that criteria “should be aids to judgement rather than rules for judgement” (Eraut, et al, 1996, page 68).

It is worth noting that the use of more holistic (indicative) criteria has sometimes been seen as the best solution to the challenge of awarding higher grades within CASLO qualifications, for instance, within GNVQs and certain BTECs (see also Newton, 2018).

Assessment criteria

Our investigation into the origins and evolution of the CASLO approach raises further questions concerning the nature and function of assessment criteria within CASLO qualifications.

Outcomes versus criteria

Particularly when criteria are presumed to function definitively, there is a risk of blurring the distinction between outcomes and their standards. That is, when having mastered a learning outcome is defined in terms of having satisfied all of its

¹⁶² This does not undermine the idea of a mastery judgement. It simply underscores the fact that mastery can only legitimately be inferred on the basis of multiple sources of relevant evidence.

associated criteria, this risks giving the impression that criteria need to be ‘mastered’ in essentially the same way as outcomes, which can lead to treating criteria as though they are little more than mini learning outcomes, as opposed to standards for judging mastery of learning outcomes.

Although, confusingly, assessment criteria do sometimes look like mini outcomes, they need to function quite differently, because their fundamental purpose is to help articulate the difference between having achieved the outcome and not (yet) having achieved it.¹⁶³ Gealy argued that this distinction became increasingly blurred over time as the CASLO approach generalised far beyond NVQs. The Further Education Unit actually defined assessment criteria as mini learning outcomes (FEU, 1995a). The QCDA provided examples of how to specify knowledge units for QCF qualifications, which looked like this:

Respond to customer requests for repairs L34

LO 1 Know the organisation’s housing stock and possible defects which require repair.

AC 1.1 Describe the types of properties which the organisation manages.

AC 1.2 Identify, using the appropriate terminology, the types of faults which can occur in these properties.

The way in which these were written renders them far more like mini outcomes than clearcut criteria, which raises the question of how assessors are supposed to identify the boundary between, say, a satisfactory description and an unsatisfactory one. This challenge becomes even more significant when knowledge and understanding requirements are abstracted from their presumed contexts of application (in the workplace) and delivered as a freestanding qualification like a Technical Certificate.

Bloom’s Taxonomy

A particularly enigmatic question concerns the role of Bloom’s Taxonomy in designing and developing CASLO qualifications. In many ways, Bloomian thinking is fundamental to the CASLO approach, as it is to many outcome-based approaches. This relates to the domain alignment goal, which identifies a need to differentiate explicitly between less complex and more complex forms of engagement with qualification content. Hence the idea of articulating what students are expected to be able to ‘do’ with the content that they are taught – recall it, analyse it, apply it, or suchlike.

¹⁶³ Or, for graded qualifications, the purpose is to help articulate the difference between operating at a lower level and operating at a higher one.

In certain CASLO contexts, Bloom's Taxonomy was used not simply as a tool for characterising learning outcome complexity but also as a tool for establishing comparable complexity across learning outcomes. This was true of the QCF, for instance, which incorporated the idea of an association between learning outcome complexity and command verb complexity, such that: 'stating' something related to content X was less complex than 'explaining' something related to content X, which was less complex than 'evaluating' something related to it. This helped to justify the idea of writing unit standards – assessment criteria – as though they were little more than mini learning outcomes. In other words, being able to 'explain' represented a certain level of competence (suitable for assessment criteria at, say, Level 2 of the QCF), while being able to 'evaluate' represented a higher level of competence (suitable for assessment criteria at, say, Level 3 of the QCF). Having said that, the use of command verbs to indicate both level-worthiness (across levels) and grade-worthiness (within levels) certainly muddied this water (see also Newton, 2018).

Ultimately, while command verbs like 'state' and 'explain' and 'evaluate' can help to scaffold qualification standards, it is important to recognise that they can do so only very roughly. For instance, as already mentioned, there is nothing in AC1.1 (above) that helps to clarify the boundary between a satisfactory description and an unsatisfactory one. Unless that gap is somehow bridged – whether by professional insight, by formal training, by detailed exemplars, or whatever – standards are likely to be applied inconsistently. Bloom's Taxonomy has no more than heuristic value, and it should not be taken too literally.

Judgements against criteria

If the foregoing analysis holds water, then writing learning outcomes (and assessment criteria) in too much detail runs a variety of risks, including:

1. increasing the assessment burden (by increasing the number of assessment judgements that need to be made)
2. decreasing the feasibility of the assessment process (by providing assessors with more judgemental scaffolding than they can deal with)
3. increasing the risk of mis-specification (by attempting to explicate the proficiency model at a level of detail that is no longer compatible with the mastery approach)

Weighing these risks against the intended benefit of disambiguation, it might well be appropriate for a qualification designer to err on the side of providing too little detail rather than too much. If so, then this would increase the need for additional mechanisms to underpin the consistent application of assessment criteria – perhaps of the sort recommended by the FEU, including exemplification materials, professional networks, and so on – especially when those criteria are written as mini learning outcomes.

Written testing

Finally, questions remain about the role of written testing within CASLO qualifications. These concern how effectively the CASLO approach can be implemented:

1. using the written test format, or
2. alongside classically designed tests (within hybrid qualifications)

In theory, there is no obstacle to implementing the CASLO approach via the written test format. It simply requires that each learning outcome is amenable to being assessed validly via a written test, and that each learning outcome has, in effect, its own mini test to enable a discrete judgement concerning mastery.

Having said that, where assessment criteria are intended to function definitively, this complicates the process by requiring a mini test for each criterion. So, the challenge is more practical than theoretical: the more learning outcomes (and assessment criteria) a unit has, the greater the burden of testing on all involved. Furthermore, if these tests were required to be taken terminally, at the end of each unit, then the consequences of failing a mini test (and therefore failing the unit) would be high. Yet, at the same time, the likelihood of failure would also be high, given the number of hurdles at which each candidate might fall. Each failed mini test would need to be resat during a subsequent session, which might actually mean having to resit full tests. So, the practical challenges are more far serious than the theoretical ones when attempting to implement the CASLO approach via written test formats, especially when they are required to be taken terminally.

Conversely, in practice, there is no obstacle to combining CASLO and classical units within a single qualification. As long as they report results in the same metric – pass, merit, or distinction, for instance – they can be combined straightforwardly. The more serious challenges to hybridisation are theoretical. For instance, if only certain units apply a mastery certification principle – while others aggregate judgements across multiple (potentially sampled) learning outcomes using a compensatory certification principle – then is that still aligned with a philosophy of mastery learning? And how should we expect qualification users to interpret results from hybrid qualifications? More pragmatically, to the extent that CASLO and classical units are designed, developed, delivered, and quality assured in quite different ways, how does hybridisation affect the financial viability of a qualification, or perhaps the quality of its implementation if resources are spread more thinly?

Lessons

We began this section by noting that it is hard to draw definitive conclusions from our analysis concerning the fitness for purpose of CASLO qualifications. After noting a couple more caveats, we will conclude with just 2 high-level lessons learnt.

Caveats

First, some of the insights that arose from our analysis were linked to, but extend far beyond, the CASLO approach. For instance, the approach played a role in making the principle of mixing and matching units (more of) a reality under the QCF. The fact that this risked undermining the coherence of study programmes – as noted by the Richard review in particular – is an important lesson to learn, but not one that bears directly on the fitness for purpose of the CASLO approach. Mixing and matching units risks undermining the coherence of study programmes regardless of how units might be designed.

Second, certain insights that appear, at first blush, to be specific to the CASLO approach are actually far more general. A major investigation into impacts associated with different modes of assessment in the learning and skills sector illustrates this (Torrance, Colley, Garratt, et al, 2005). On the one hand, the authors concluded that the move towards criterion-referenced assessment and competence-based assessment had increased uptake and improved learner achievement and progression. However, on the other hand, the very transparency that was fundamental to these approaches had also encouraged instrumentalism, thereby reducing the challenge of learning and the validity of qualification results:

We have identified a move from what we characterise as assessment *of* learning, through the currently popular idea of assessment *for* learning, to *assessment as learning*, where assessment procedures and practices may come completely to dominate the learning experience, and ‘criteria compliance’ comes to replace ‘learning’. This is the most significant challenge confronting assessment in the LSS: balancing the explicitness of learning objectives and instructional processes against the validity and worthwhileness of learning outcomes.

(Torrance, Colley, Garratt, et al, 2005, page 2)

Remember that exactly this concern was identified in the CAVTL report, and then in the Whitehead report. In terms of this second caveat, however, the most important point to note from the Torrance report was that they expressed exactly the same concerns regarding learners who were studying AVCE and A level programmes. In other words, criteria-focused coaching and the concomitant risk of criteria compliance was proving to be similarly problematic for post-millennium classically designed qualifications.

Putting caveats of this sort to one side, we believe that our investigation into the origins and evolution of the CASLO approach enables us to reach 2 major conclusions related to fitness for purpose:

1. the CASLO approach cannot be said to be universally fit, nor universally unfit, for purpose
2. it is not easy to render CASLO qualifications fit for purpose, and in some ways they are harder to render fit for purpose than classically designed qualifications

Neither universally fit nor unfit

On the one hand, our analysis has provided insufficient reason to conclude that the CASLO approach, or outcome-based approaches more generally, are fundamentally flawed and therefore universally unfit for purpose.

Particularly in recent years, there has been a tendency for scholars to treat the NVQ model as though it reflected the essence of all outcome-based approaches to qualification design: “a ‘pure’ learning-outcomes approach” (Young & Allais, 2009, page 3), or “a ‘full-blooded’ version of the outcomes-based design philosophy” (Winch, 2020, page 169). The implication is that, by striking at the heart of the NVQ model, we strike at the heart of the outcome-based approach, per se, and therefore at the heart of the CASLO approach too. On the basis of an analysis of this sort, Winch recently declared that: “outcomes-based qualifications are not fit for purpose” (Winch, 2023, page 20).

Our own analysis does not lead to such an extreme conclusion, at least, not as we define the essence of an outcome-based qualification.¹⁶⁴ More specifically, we are not convinced that it is legitimate to generalise from criticisms of NVQ theory and practice – both of which involved features that were highly idiosyncratic – to outcome-based qualifications more generally. Furthermore, our analysis reveals that there is a positive as well as a negative story to be told concerning the introduction of the NVQ system in England, which is also true in respect of CASLO qualifications more generally, and of outcome-based qualifications more generally than that.

On the other hand, our analysis has provided sufficient reason to conclude that the CASLO approach – and outcome-based approaches more generally – are not universally fit for purpose. For instance, we believe that it was a mistake to assume that all qualifications can straightforwardly be shoehorned into this design template, as those who designed the QCF appear to have assumed. There will be contexts and purposes that make adoption of the CASLO approach more challenging, such

¹⁶⁴ We have had many fruitful conversations with Chris Winch over the different ways in which outcome-based approaches are defined in different contexts, and how unhelpful this can be. His definition seems to be somewhat narrower than ours.

as when results are used for purposes that have high stakes for teacher-assessors (although it is fair to say that classically designed qualifications may also be less fit for purpose under such circumstances). There will also be cohorts for whom adoption of the CASLO approach is likely to be less beneficial. For example, the benefits of implementing a mastery learning philosophy might be lower for qualifications that tend only to attract highly motivated and highly capable students.

The decision to adopt the CASLO approach ought to follow from a rigorous analysis of the purposes that a particular qualification (or type of qualification) is intended to serve, given the nature of the cohort for whom it is being designed, and the contexts within which it will need to operate. Sometimes an analysis of this sort will recommend an outcome-based approach, such as the CASLO approach, but not always.

Challenging to ensure fitness for purpose

One part of the explanation for why it is not easy to render CASLO qualifications fit for purpose relates to the fact that the CASLO approach has tended to be used for TVET qualifications, and it is harder to design qualifications to be fit for TVET contexts than for traditional academic ones.

Traditional academic qualifications benefit from their grounding in disciplines that have accumulated wisdom, status, and rigour over decades if not centuries. They are overseen by scholarly communities of practice, which include established teaching and learning communities, and they typically involve large numbers of students and well-qualified staff. The same cannot be said for many TVET qualifications. TVET qualifications also tend to attract students who have been less successful on prior academic qualifications, including students who have straightforwardly been failed by them (in both senses of the word). Finally, whereas many of the outcomes associated with academic qualifications are readily amenable to the written exam format, this is often not the case for TVET qualifications. Alison Wolf put it like this:

However, in drawing a line under the history of NCVQ, it is important to remember what first inspired the 'competence' movement and the advocacy of portfolio based techniques. In most vocational contexts, it is virtually impossible to develop timed, paper based, externally marked assessments which carry conviction as valid samples of the skills and competencies concerned. Research evidence (summarised in Wolf, 1995a) confirms that such tests generally have very limited value as predictors of workplace performance. Legitimate concerns with the reliability and manageability of assessments do not, in themselves, help us to develop techniques which also generate valid judgements about capabilities and performance. In Britain, the failures of NVQ and GNVQ policies leave these problems unsolved but no less pressing than before.

(Wolf, 1998, page 442)

Another part of the explanation for why it is not easy to render CASLO qualifications fit for purpose stems from the fact that outcome-based qualifications attempt to provide far more structure than classical ones. This additional structure is intended to help scaffold teaching, learning, and assessment. Yet, it generates risks too.

The present report highlights the difficulty of designing an outcome-based qualification model that is right first time. Significant difficulties were encountered with all of the earliest innovations, including the TEC and the BEC models, the NVQ model, the GNVQ model, and so on. Each of these models was refined and reformed over time.

Of course, all qualification models are likely to be refined and reformed over time. Yet, outcome-based qualifications can be particularly difficult to get right because, by design, they model outcomes and criteria far more explicitly and in far more detail than classical ones. This increases the risk of modelling them inadequately or inappropriately. This includes the risk of specifying outcomes (and criteria) in too much or too little detail. It also includes the risk of pitching standards at a level that is either too demanding or insufficiently demanding. These challenges can be harder for school-based qualifications than for work-based ones.¹⁶⁵

These challenges are exacerbated by the fact that outcome-based qualifications are less forgiving of inadequate or inappropriate design decisions. If a classically designed qualification specified too much content, then this would no doubt place students and teachers under unnecessary pressure as they studied for their exams. Yet, because exam standards are ultimately applied post hoc, adverse consequences can be mitigated by setting grade boundaries lower than originally anticipated. In contrast, outcome-based qualifications specify required outcomes and standards in advance, which makes it hard to mitigate adverse consequences arising from having specified too many, or unduly demanding, outcomes.

These challenges emphasise the critical importance of piloting and trialling outcome-based qualifications. The broader the scope of the reform process – whether at the level of a single qualification, or a qualification type, or a qualification framework – the more extensively the outcome-based model will need to be piloted and trialled. These challenges also underline the level of professional insight and expertise that

¹⁶⁵ Where qualification design is driven primarily by the need to certify occupational competence, certain of these factors (such as qualification demand) will not be negotiable, which may necessitate holding other factors open (such as qualification time). Where qualification design is simply driven by the need to provide a suitable course of learning for a particular cohort of students who will be studying for a set period of time, the challenge of pitching outcomes and standards effectively will loom large.

will be required of anyone responsible for designing and developing outcome-based models and qualifications. We will return to this point shortly.

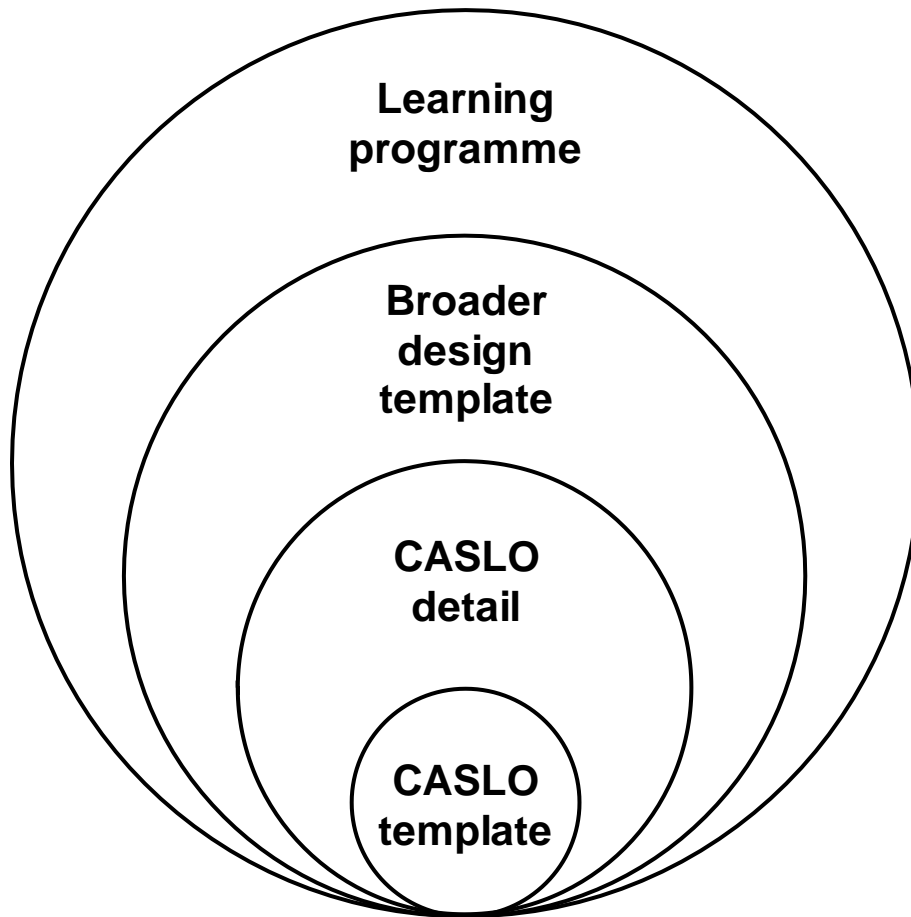


Figure 17. Securing the success of a CASLO qualification

The CASLO template

Perhaps the most important lesson to learn is that it is hard to say anything definitive about the CASLO approach, per se, because the validity of any particular CASLO qualification, as well as the nature of its educational and societal impacts, will depend upon how the basic CASLO template is brought to life. Ultimately, the CASLO approach is little more than a high-level design template that provides a foundation for building a broader qualification design template, which is bespoke to the qualification in question. It is the adequacy and appropriateness of this broader qualification design template that really matters (as well as how it gets put into practice, of course). Figure 17 attempts to illustrate this idea, indicating that multiple

layers of planning and decision making are required to secure the success of a CASLO qualification.

Qualification design always needs to begin with a rigorous analysis of the purposes that a new qualification (or qualification type) will need to serve, given the nature of the cohort for whom it is being designed, and the contexts within which it will need to operate (Newton, 2023). Let us assume that an analysis of this sort recommends adopting the CASLO approach. Ultimately, all this means is that our proficiency model needs to be specified in terms of learning outcomes, assessment criteria, and a stringently applied mastery principle. This is the 'CASLO template' level of analysis from Figure 17. Upon this foundation, CASLO qualification design might diverge in all sorts of different ways, and these subsequent design decisions will be critical to the success, or otherwise, of the qualification in question.

At the 'CASLO detail' level of analysis, we will need to decide exactly how our learning outcomes and assessment criteria are to be represented. This is far from a trivial decision, as the NVQ story taught us: the approach that was advocated by the NCVQ – the NVQ competence model – proved to be highly controversial. The subsequent QCF approach was far more traditional, tending to frame learning outcomes as elements of knowledge, skill, or understanding. QCF assessment criteria were not unproblematic, though, often having been written more like mini learning outcomes than as definitive criteria. TEC and BEC designers also experimented with different approaches to representing outcomes and criteria.

Note that the decision concerning how best to represent learning outcomes and assessment criteria is just the first step at this level of analysis. Subsequent steps include determining an appropriate grain size for writing learning outcomes, deciding how the assessment criteria ought to function and how many to include, determining how to build depth into outcomes and criteria (for example, via Bloom's Taxonomy), and so on.

The next level of analysis concerns the 'broader design template' and this is where all of the remaining qualification design decisions get made. On the one hand, this means fleshing out details of the assessment procedure that underpins the qualification. On the other hand, it means explicating implications for teaching and learning, particularly when a specific approach is anticipated, which might also affect the design of the assessment procedure. Again, decisions at this level will be critical to the success of the qualification.¹⁶⁶ Obviously, the mastery principle will be relevant

¹⁶⁶ This, incidentally, explains why some awarding organisations were extremely uncomfortable with the idea of building the Qualifications and Credit Framework on the foundation of a bank of shared units. It left open the possibility that an inexperienced awarding organisation might build suboptimal procedures and approaches on the foundation provided by those shared units, ultimately undermining

here, as mastery learning is fundamental to mastery certification. However, there might be broader issues to consider, for instance, the extent to which the qualification promotes a problem-based approach to learning. Considerations at this level of analysis explain why the CASLO family is so broad, embracing both NVQs (at one end of the spectrum) and BTECs (at the other).

This brings us to the 'learning programme' level of analysis, which is where qualification requirements get translated into learning journeys and learning experiences. Ultimately, this is the responsibility of teachers and trainers in colleges and workplaces. Yet, there is plenty of scope for awarding organisations to provide support and guidance at this level too. More to the point, qualifications need to be designed with this stage in mind, so as not to incorporate features or processes that might frustrate the production of effective learning programmes. This is why Stanton has consistently argued that teachers and trainers need to feed into qualification design decisions, both from the outset and throughout the process. We will return to this level of analysis later under the 'whole system reform' heading.

Reforming TVET qualifications

The story of the CASLO approach comprises a very large part of the story of regulated TVET qualifications in England over the past half century, spanning multiple major reforms. So, we end this report by considering lessons that we might be able to learn about TVET qualification reform more generally, based on experiences from designing and implementing CASLO qualifications in England.

One of the key challenges that we have faced so far (in this report) has been to distinguish goals that were specific to the CASLO approach from far broader ones (that had little or no bearing on why the approach was adopted). However, we are now going to begin this final section by focusing specifically on one of most important of the broader TVET reform goals, namely rationalisation.

Rationalisation

Rationalisation through increasingly explicit regulation has been the dominant meta-narrative to stories concerning qualifications in England over the past century. This has involved creating systems, frameworks, and agencies, where none previously existed, and attempting to reduce the number and variety of qualifications on offer. This has been true of all types of pre-university qualifications. For instance, the Acland report on 'Examinations in Schools' (Acland, 1911) led to the first national exam system in England, based upon the School Certificate and Higher School

their credibility (and the credibility of the qualifications that they were included within). Unfortunately, this did happen.

Certificate. These exams were administered by a small number of universities and overseen by the Secondary Schools Examinations Council, a new co-ordinating body that reported to the Board of Education.

Although a number of TVET qualification subsystems were established during the early years of the 20th century, including the system of Ordinary and Higher Nationals – and although major players like the City & Guilds and the RSA had their own organisation-specific qualification systems – the general lack of structure and co-ordination became increasingly poignant over time as problems became increasingly apparent. Government was resistant to exercising control over the TVET landscape, seeing training principally as a matter for industry and commerce. Yet, by the end of the 1950s, it was clear that this hands-off approach was not working, and strong central intervention became part of the solution, initially in relation to training and subsequently in relation to qualifications.

The TEC and the BEC were established to rationalise the technician qualification subsystem during the early 1970s. By the end of the 1980s, the NCVQ had been established to rationalise the entire landscape of TVET qualifications. The NCVQ was the first agency to introduce a national qualification framework, to which all qualifications would be accredited. This was a powerful rationalising tool, as it was designed to ensure that all qualifications adopted exactly the same high-level design template, thereby attempting to prohibit the variety of approaches that had become a hallmark of TVET qualification provision prior to that. Since the 1980s, government-sponsored agencies have controlled the TVET qualification landscape using a succession of qualification frameworks. Some of these have been more prescriptive (for example, the Qualifications and Credit Framework) and others less so (for example, the current Regulated Qualifications Framework).

Fewer qualifications

Consistently, over time, the goal of reducing the number of qualifications within the TVET landscape has been pursued in the name of creating a simpler system. The idea of counterproductive complexity was captured in the idea of there being a qualification ‘jungle’ that needed to be cleared, leaving only the finest specimens behind. Clearing the so-called ‘jungle’ in this manner would make it more navigable for learners and for employers – and for teachers, trainers, parents, careers advisors, and other stakeholders too – increasing the likelihood that learners would locate high value qualifications (and progression routes) best suited to their needs and aspirations.

Unfortunately, experience has taught us that it is harder than it sounds to reduce the number of qualifications in the market. Indeed, reforms intended to rationalise the system can result in a proliferation of new qualifications, which is a phenomenon that

was recognised as early as the 1960s (Haslegrave, for instance, identified a proliferation of new courses in the wake of the 1961 white paper).

Prescriptive qualification frameworks that formally introduce new qualification types, such as the NVQ framework and the QCF, appear to be particularly vulnerable to the risk of qualification proliferation. When the NVQ framework was introduced, the NCVQ anticipated that all regulated TVET qualifications would be subsumed within it. Yet, many existing qualifications continued to exist beyond the framework, and even the new GNVQ was not actually accredited to the NVQ framework. This necessitated a broader framework, and the National Qualifications Framework was introduced in 2000.

Half a decade after NVQs had been introduced, the FEFC observed that there was “little evidence” that this had “led to a rationalisation of qualifications offered within colleges” (FEFC, 1994a, page 12). Indeed, Williams concluded that the introduction of NVQs had “simply added to the existing jungle” (Williams, 1999, page 218).

A similar thing happened when the QCF framework was introduced, which was also originally anticipated to subsume all regulated qualifications. A larger proportion of the qualification market ended up being regulated under the QCF than under the NVQ framework. Yet, plenty of TVET qualifications continued to exist beyond it, regulated under the NQF. Furthermore, just as under the NVQ framework, qualifications actually proliferated under the QCF (Oates, 2013a).¹⁶⁷

Part of the challenge in exercising control over the system via qualification frameworks is that the state, in England, has only limited powers, given that qualifications are provided by independent awarding organisations. Government can attempt to persuade these organisations to replace their existing qualifications with new ones – and can heavily influence this via funding and performance table requirements that impact on qualification consumers – but it has tended not to go so far as to prohibit awarding organisations from continuing to offer old qualifications (alongside new ones). Where the old qualifications have continued to be sufficiently attractive to a sufficiently large number of consumers, the awarding organisations have had an incentive to continue offering them. Bear in mind that even school and college qualifications sometimes have national or international markets that are not necessarily swayed by funding or performance table pressures.

Furthermore, on a number of occasions, government has capitulated to pressure to retain certain qualifications, ultimately recognising that they do fulfil an important need (for certain students and qualification users) that was not adequately fulfilled by

¹⁶⁷ The significance of this proliferation is hard to judge, as part of the rationale for the QCF was to accommodate awards that would not previously have been accredited (under the NQF). This shines a light on the elephant in this particular room, which is the plethora of qualifications that have always existed beyond the regulated qualifications market (see PricewaterhouseCoopers, 2005).

their notional replacements. This happened with the NVQ framework, when funding was approved for certain existing qualifications that failed to satisfy the NVQ accreditation criteria. It also happened when successive GNVQ reforms ran out of steam, and pressure to withdraw BTECs was lifted.

The BTEC example is also interesting for highlighting the importance of branding. Even qualification users (like employers) rarely understand how qualification systems operate, beyond familiarity with the qualifications that they (as individuals) took when they were at school and college, and familiarity with qualification suites that have traditionally underpinned their own sector. For instance, it was widely observed how poorly employers understood the NVQ system, and they proved to be particularly resistant to being educated. In a context like this, the maxim 'better the devil you know' is likely to hold true. The BTEC council was therefore very shrewd in retaining the 'National' and 'Higher National' nomenclature following withdrawal of the old system of Ordinary and Higher technician awards. And, when the Council was ultimately disbanded, continuation of the BTEC brand was equally shrewd.

One final observation concerns a more subtle risk of qualification proliferation. New prescriptive frameworks tend to act like 'new brooms' that inadvertently open the market up to new providers. This occurs when a new framework aims to replace existing qualifications with a new type of qualification, like an NVQ or a QCF qualification, which is built to comply with new design rules. Where a single provider might previously have won market share for an established qualification over an extended period of time, new design rules can wipe the slate clean, levelling the playing field for the new awards. The replacement qualification might now be offered by multiple awarding organisations.

Qualifications certainly proliferated under the QCF as identical units were packaged into the same qualification offered simultaneously by multiple awarding organisations. Beyond proliferation, per se, this approach ran the risk that some of these new qualifications would be developed and delivered far less effectively than others. This was a problem under the NVQ framework (Raggatt & Williams, 1999) and proved to be so under the QCF. Indeed, one of the principal reasons for withdrawing the QCF was that it had (intentionally) opened the market up to new, inexperienced awarding organisations, some of which developed qualifications that managed to satisfy QCF accreditation criteria whilst ultimately lacking validity (Ofqual, 2014b).

Homogenised qualifications

Consistently, over time, the goal of homogenising the design of qualifications within the TVET landscape has been pursued in the name of creating a stronger system. Instead of leaving the finest specimens to flourish, post-deforestation, government and its agencies in England have tended to want to cut the entire jungle down and

replant it as an arboretum of perfectly cultivated trees. This was the logic of the NVQ framework, which specified accreditation criteria in such detail that it effectively called for universal qualification reform, and it was the logic of the QCF too.¹⁶⁸

The rationale for starting from scratch with a brand new qualification model appears plausible and attractive. In the NVQ context, it resulted from having identified serious problems with existing qualifications that appeared to be linked to specific design features. By specifying a different approach to qualification design, these problems were intended to be addressed. This rationale directly explains the origins of the CASLO approach in England.

However, while plausible and attractive, the desire to homogenise qualifications – which is typically characterised as ‘strengthening’ them – runs multiple risks, which relate to the possibility that these newly specified design features might:

1. not actually solve the problems that they were supposed to solve
2. introduce new unanticipated problems
3. work well for certain qualifications but not for others
4. prevent the use of even better design features (thereby also stifling innovation and competition)

Successive reforms to qualification systems and frameworks in England have swung the pendulum of change a very long way away from the situation that pertained in the 1950s, when the leading TVET awarding organisation was proud to declare that there was no such thing as a typical City & Guilds qualification. Back then, responsiveness to local needs and aspirations was the guiding principle. Even when the TEC and the BEC took steps to rationalise the technician qualification subsystem, during the 1970s, they still embraced the idea of responsiveness by implementing a validation model, whereby centres were responsible for programme development and were encouraged to customise provision accordingly, albeit within parameters. Having said that, even during the 1970s and 1980s, the demand for customisation of this sort was not overwhelming, and centres often preferred to rely on standard units (which, it is fair to say, still offered a certain amount of flexibility). Nowadays, customisation of BTEC units occurs only rarely, by exception.

Interestingly, CASLO qualifications have tended to be homogenised in a manner that retains a certain amount of responsiveness. On the one hand, their core characteristics mean that all CASLO qualifications are highly constrained by the detail in which learning outcomes and assessment criteria are specified, and by specifying the mastery requirement. On the other hand, when those learning

¹⁶⁸ Stanton, Morris, & Norrington (2015) identified a failure to learn lessons from the past – alongside the clear preference of inexperienced policy makers for a ‘clean sheet’ approach to problem solving – as 1 of 4 main reasons why agency-designed qualifications have failed in the past.

outcomes are specified at a fairly high level, this can facilitate a certain amount of customisation, which was encapsulated in the goal of (minimal) cross-context domain personalisation. This can also be facilitated by incorporating optional units, or routes, which also helps to improve responsiveness while keeping proliferation in check (as emphasised by both the Beaumont and CAVTL reports). Likewise, by exploiting the transparency inherent in the approach, this permits a certain amount of operational flexibility (which was encapsulated in the goal of qualification efficiency).

Yet, qualification homogenisation is no panacea or silver bullet, even when cunningly operationalised. For instance, centrally specified NVQ standards proved to be problematic when defined either too narrowly, where learners failed to acquire competencies that they actually needed, or too broadly, where learners were required to acquire competencies that they did not actually need. Although each NVQ standard was intended to represent a consensus position, which established a certain amount of democratic legitimacy, this also meant that they were vulnerable to not reflecting the needs of any particular employer very well, or to only representing the needs of certain employers when the consensus building was unrepresentative (see Oates, 2004, on NVQ standards as pragmatic constructs).

In certain areas of certain sectors, problems of this sort challenged the very idea of a national qualification. The review of NVQ assessment produced by Eraut, et al (1996) brought this challenge home. To some extent, a qualification framework is a useful fiction, which helps to make sense of similarities across otherwise disparate contexts, but that cannot be taken too literally. For instance, just because 2 NVQs were both classified at Level 2, that did not literally mean that they had exactly the same level of demand and required exactly the same amount of training time – particularly not under the NVQ framework where standards were designed to be role specific. Thus, the Eraut report explained how carpentry and joinery were much more demanding than painting and decorating, which needed to be recognised in terms of both training time and funding. The report went even further by arguing that:

It is impossible to have a national system of qualifications based on current competence at work; because we do not have a national system of working practice. But it is possible to have a national system of competence-based units which are not mandatory for all trainees.

(Eraut, et al, 1996, page 67)

The Wolf report reached a similar conclusion:

However, because a complex modern economy has a correspondingly complex occupational structure, central attempts to impose a neat, uniform and 'logical' structure on it always fail.

(Wolf, 2011, page 57)

It seems likely that the 'sweet spot' lies somewhere between the extremes of, on the one hand, qualification design being determined by the accreditation requirements of a single overarching framework and, on the other hand, of qualification design being entirely bespoke. Exactly where on that continuum the sweet spot lies is unclear. Indeed, it might well be in a different location for different qualifications (see also Oates, 2013b).

Clearly, though, the higher the level at which a qualification framework is implemented – a single overarching framework being the highest level – the greater the risk of it failing to accommodate the diversity of the qualifications that it subsumes. This underlines the necessity of comprehensive trialling and piloting: the higher the level of the framework, the more extensive this trialling and piloting will need to be. Yet, the very complexity of occupational structure in England renders trialling and piloting (and, subsequently, evaluating) extremely challenging. Trials undertaken in one area of one sector are quite likely not to generalise to another.

In theory, the decision to implement one qualification model as opposed to another – for instance, the CASLO approach rather than a classical one – ought to be taken on the basis of a detailed analysis of the purposes and cohorts that it is intended to serve, and the contexts within which it is intended to operate. If, instead, this decision is taken (by fiat) at the qualification framework level – by building it into accreditation criteria – then the framework designer will need to be correct in assuming that the model will not be inappropriate for any of the qualifications subsumed by the framework. Both the NVQ framework and the QCF fell short in this respect.

One final homogenising tendency is worthy of mention, which involves making TVET qualifications look more like classical ones, typically in the name of strengthening them, but sometimes also in the name of improving parity of esteem. This happened repeatedly to GNVQs. By the time the AVCE model was introduced, in 2000, the qualification had effectively lost its way. Shorn of its original defining characteristics, it no longer seemed to be fit for any particular purpose, and it was soon withdrawn. The qualification that followed in its footsteps was literally an A level, the Applied A level, although it was not a great success.¹⁶⁹

School and college performance table qualifications were partially homogenised by DfE requirements that followed in the wake of recommendations from the Wolf report. This led to the hybridisation of many CASLO qualifications, whereby the changes effectively enforced a classical approach on certain units, while other units continued to operate under the CASLO approach. The impacts of hybridisation along these lines are still to be thoroughly investigated.

¹⁶⁹ They were more popular as AS awards than as full A levels, but uptake of both qualifications declined steadily over time until they were withdrawn (Sutch, Zanini, & Benton, 2015).

Control

As government has increasingly intervened to rationalise qualification systems in England, it has increasingly exercised control over those systems through networks of co-ordinating and regulatory agencies, of which Ofqual is one. This raises important questions concerning where responsibilities for different aspects of qualification design, development, delivery, and review ought to reside, and what the consequences of different configurations might be.

These questions are particularly challenging because of how widely shared these responsibilities have traditionally been, owing to the fact that qualifications operate at the interface between curriculum, pedagogy, and assessment (that is, at the interface between quite different stakeholder groups). This is further complicated within the TVET landscape by the disparate needs of employers, not to mention ambiguity over who ought to be responsible for meeting those needs, let alone how to ensure sufficient engagement within a largely voluntaristic system.

Central intervention

The assumption that employers ought to take responsibility for their own training needs acted as a brake on central intervention until the early 1960s. The establishment of a Central Training Council and Industrial Training Boards began to change this. Initially, the focus was upon training, and training standards. But attention soon turned to qualifications, and the certification of competence, when the TEC, the BEC, and the NCVQ were established. Through the NVQ system, government simultaneously exercised control over standards for education and training (National Occupational Standards) and methods for assessing and certifying those standards (National Vocational Qualifications).

Critically, the CASLO approach was part of this initiative. The features that comprise it were implemented fully within NVQ awards, and partially within TEC and BEC awards. The existing awarding organisations were already largely signed up to the general idea of specifying qualifications in terms of learning outcomes and assessment criteria, but it was government-sponsored agencies that implemented this approach at scale.

Responsibilities

Because qualifications operate at the interface between curriculum, pedagogy, and assessment, there will always be some debate over how best to achieve their co-ordination. For over a century, awarding organisations in England have played a central co-ordinating role, and were assumed to be where the buck stopped in terms of accountability for the quality of those qualifications. Within the TVET landscape,

the pre-eminent awarding organisation during the first half of the 20th century was City & Guilds. It relied heavily upon advisory committees for guidance on preparing course syllabuses, ensuring that the views of educationists, employers, and trade unions were all heard. For the original Ordinary National and Higher National awards, this co-ordinating role was assumed by Joint Committees representing the relevant professional institution(s), the Ministry of Education, and teacher associations.

Although central co-ordination is clearly critical, and although there are good reasons for government to be part of that co-ordinated activity, experience also teaches us that there are challenges associated with increasingly centralised control. Indeed, the Wolf report argued against government control, on the basis that having to assume ultimate responsibility will make it “correspondingly impossible to be honest” when qualification reforms go wrong (Wolf, 2011, page 9).

This raises an important question concerning ownership of, and responsibility for, the CASLO approach. If a qualification model, like the CASLO approach, is specified as a qualification framework requirement, then there is a lot of sense in it being ‘owned’ centrally by government. Yet, it is the responsibility of an awarding organisation to design, develop, and deliver qualifications using this high-level design template. This makes it unclear exactly where the buck stops in terms of the quality of qualifications within the framework, particularly where it is impossible to disentangle issues arising from the centrally specified model and issues arising from features and processes specified by an awarding organisation. In the worst case scenario, it disposes awarding organisations to abrogate responsibility for quality, instances of which appear to have occurred under the QCF, when it was entirely unclear where responsibility for the quality of units was located.

Expertise

Increasing central control also affects the structure of knowledge within a sector, in this instance the qualifications sector. While the awarding organisations were experimenting with outcome-based approaches during the 1970s and 1980s, the NCVQ was to become the powerhouse of the CASLO approach, with an increasingly powerful internal research function (Ecclestone, 1998). With the Training Agency, it oversaw the design and implementation of the NOS-NVQ model. Based on functional analysis, this model was technically demanding to implement. It was an expert system that required expert practitioners, who were few and far between even then, and the system relied heavily upon centrally contracted consultants.¹⁷⁰

¹⁷⁰ Raggatt & Williams observed that these technical consultants were alleged to be the only ones who truly understood functional analysis (Raggatt & Williams, 1999). The credibility of the

Perhaps the biggest risk associated with locating critical knowledge centres within government-sponsored agencies relates to what happens when their functions are dissolved and they are disbanded. The agencies that followed the NCVQ and the Training Agency did not 'own' the CASLO model in quite the same way or to the same extent. Furthermore, because the awarding organisation sector increasingly developed qualifications on the back of NOS, rather than developing standards from scratch, they increasingly became users rather than owners of standards. In essentially the same way, they increasingly became users of the CASLO approach rather than owning it. This may be at least part of the reason why there is so little published research into the approach. If the awarding organisations were not ultimately responsible for it, then were they similarly absolved of responsibility for researching and evaluating it?

So, who are the CASLO experts today, and in what kind of organisation might we expect to find them? It is not entirely clear. The fact that the present programme of research was even necessary raises a fundamental question concerning the structure of knowledge within the qualifications sector. One thing is for sure, though: however that knowledge structure is configured – and it could legitimately be configured in a variety of different ways – it must always be the case that whichever organisation is ultimately responsible for qualification, or qualification system, design decisions, it must be equipped with the necessary expertise to make those decisions effectively.

As far as expertise is concerned, the qualifications sector faces an unfortunate, if not ironic, impediment: there is no formal 'certification profession' in England and there are no generally accepted certification qualifications. Unlike professions such as medicine, or accountancy, you do not qualify to become a certification professional, you tend to fall into the qualifications sector, and accumulate relevant knowledge, skills, and understanding on the job. At the very least, this lack of formal recognition of the expertise required to design and implement certification systems is odd. At worst, it constitutes a systemic threat to the quality of professional practice in the qualifications sector in England.

methodology was also heavily dependent upon which particular industry experts the technical consultants enlisted. Lester (2015) noted that functional analysis was particularly vulnerable to relying on senior members of relevant occupations who lacked insight into current working practices or contexts. Parkes (1994) noted the credibility threat that arose when standards were developed in collaboration with sectors that lacked a strong tradition of commitment to, and investment in, training. Callender (1992) identified risks associated with relying upon experts within industries that were neither co-ordinated nor co-operating in their contributions.

Pragmatics

Finally, our analysis provides insights into how TVET qualification reform has been undertaken over the years, including lessons concerning how reforms can go wrong.

Speed

The introduction of Curriculum 2000 is remembered less for the new AVCE, which was to fizzle out within a few years, and more for the first new modular A level awards, which exploded with a bang shortly after A level results day in 2002 (Tattersall, 2007). Anomalously low results in certain units led to multiple official inquiries and ultimately to 18 unit grade boundaries being lowered and to 9,800 unit grades being raised. It also led to the Chairman of the QCA being sacked, and to the Education Secretary resigning shortly afterwards. A Select Committee on Education and Skills report into the crisis concluded:

On the evidence presented to us, we conclude that the events of last Summer were not caused by the manipulation of the examination system but by confusion arising from the introduction of the A2 exam without adequate trials.

(HCESC, 2003, paragraph 78)

The government responded by acknowledging:

We recognise absolutely that there are lessons to be learnt for the future about the way in which we implement major reforms of this sort. Detailed planning and extensive trialling is essential so that we can be confident that all systems are in place and that teachers and examiners are fully trained in new requirements before they are introduced.

(DfES, 2003b, paragraph 5)

Qualification reforms are always under pressure to deliver results as soon as possible. Even when piloting and trialling activities do take place, there is no guarantee that problems with the new systems and qualifications will be spotted and resolved. And, of course, it takes a long time to pilot and trial a qualification.¹⁷¹ Both the QCF and the AVCE were at least notionally 'trialled' before going live, yet, they both ultimately failed.¹⁷²

¹⁷¹ New syllabuses take a long time to be agreed. They need to be disseminated significantly before first teaching, and the new courses will run for a year or two before the first awards are made.

¹⁷² Arguably, these might be better described as 'phased introductions' rather than as genuine 'trials' given the widespread presumption of system change.

What seems eminently clear, however, is that systems and qualifications as complex and novel as the NVQ system, or the GNVQ, stand little hope of succeeding without extensive piloting and trialling over a significant period of time. Scaling up too soon can cause serious reputational damage to the qualification brand when things go wrong on a national scale. In this situation, even if the qualification model is rapidly revised to address these issues, the initial reputational damage may be sufficient to have undermined the credibility of the new qualification brand (and, for qualifications, credibility is the bottom line, more so even than validity). In the words of Bent Flyvbjerg and Dan Gardner – think slow, act fast. It is far better to make mistakes in the planning stage than during delivery (Flyvbjerg & Gardner, 2023).

Given their complexity and novelty, it seems fair to conclude that both NVQs and GNVQs were scaled up to national rollout too quickly, with shambolic consequences. The fact that they were both based upon a new qualification model – not to mention adopting other innovative and idiosyncratic features and processes specific to each of them – greatly magnified the risks associated with rushing their implementation. Critically, these reforms did not simply have radical implications for assessment, they also had radical implications for teaching and learning, which is a point that we will return to shortly.

Engagement

These radical reforms also provide insight into the importance of engaging stakeholders effectively. The introduction of NVQs provides an important case study in its own right, given just how radical the reform was, and just how extreme the critique from certain quarters proved to be. It seems reasonable to conclude that the NCVQ should have adopted a better strategy for engaging stakeholders, although communications would always have been highly challenging to manage bearing in mind that the reform was explicitly intended to disrupt existing practices.

Unfortunately, the NCVQ shook the system so violently, and rapidly, that it ended up metaphorically at war with many academic educationists. Although NCVQ research teams did actively collaborate with academics who generally supported their approach (such as John Burke), and with those who were prepared to contribute as critical friends (such as Alison Wolf), they struggled with those who were more openly hostile (such as Alan Smithers). As researchers surveying the scene 2 to 3 decades later, we see evidence of these rifts even to the present day.

Employer engagement is a perennial challenge for TVET in England, for many varied and complicated reasons (see, for example, Stanton & Bailey, 2004; Unwin, et al, 2004; BIS & DfE, 2016; Huddleston, 2020). The basic challenge is how to ensure that the voice of employers is sufficiently loud, representative, and clearly articulated to ensure that their needs are adequately reflected in TVET qualification design. However, to the extent that employers are also expected to play an active role in

qualification delivery – the idea at the heart of the NVQ model – the challenges multiply significantly.

Inevitably, the more we expect employers to be actively involved in the technical details of qualification design and delivery, the more accessible those technical details will need to be. This runs the risk of creating tension between the need for qualification rigour and the need to be responsive to employer demands for simple systems and processes (as discussed above). When employer engagement is required, yet can only be incentivised and not enforced, this increases the risk of capitulating to demands that result in oversimplified systems and processes.

Once again, the challenge is how to balance the scales, such that employers are sufficiently engaged, but not so much so that they end up with responsibilities that qualification specialists could deliver more effectively.

Overstepping the mark

One of the most important risks highlighted by our research is that reforms overstep the mark. This can happen when things go wrong, but the proposed solution swings the pendulum of change too far in the opposite direction. When this happens, it is likely that the new system will fail for exactly the same reason as the old one, albeit operating in reverse.

Once again, the NVQ model provides a good illustration of this phenomenon. Previous qualifications had been tailored to off-the-job education and training, and were rightly criticised for focusing too much on theory and book knowledge. Yet, the solution – the NVQ model – ended up being tailored to on-the-job training, and was rightly criticised for focusing too much on practice and informal learning. This threat was identified early on (see Black & Wolf, 1990) but the model was already too entrenched by then for this pendulum swing to be reined in effectively. This underscores the risks associated with reactive, knee-jerk reform (Lum, 2015). One final consideration is that the more radical the reform in question, the more challenging it becomes to implement, which increases the importance of not overstepping the mark with a reform that is more radical than it actually needs to be.

The QCF provides a slightly different illustration of overstepping the mark, which involved generalising to all learners a solution that seemed to work well in the context of adult learners. The utility of a flexible unit-based framework had been demonstrated in the OCN context: effectively empowering adult learners – especially women who wished to return to the workforce – to re-engage with the discipline of learning, one small step at a time. However, when generalised to the QCF context, the idea that all learners would benefit from being able to mix and match units to form bespoke qualifications was never substantiated. Indeed, the flexibility built into the QCF had the unintended consequence of facilitating, if not promoting, incoherent curriculum programmes.

In addition to overgeneralising a solution to the genuine problems faced by returning adult learners, the QCF also engineered solutions to problems that proved not to be significant after all. Paramount, in this respect, was the facility for transferring credit across awarding organisations, for which there proved to be very little demand indeed (Ofqual, 2010b; 2011b).

Whole system reform

By far the most important risk highlighted by our research is that qualification reforms are conceptualised and operationalised too narrowly, with insufficient attention to the wider education and training changes that are necessary for those reforms to bed in (see also CAVTL, 2013; Oates, 2004; 2013a; 2013b). Qualification reforms are best understood as education and training reforms that are initiated through changes to certification requirements. When considered from this perspective, the importance of adequately involving and supporting teachers, trainers, centres, and learners right from the outset becomes harder to overlook.

In terms of involving key protagonists, Stanton (2012) argued that standards need to be developed in collaboration with teachers and trainers, to maximise the likelihood of their being meaningful and useful when subsequently translated into learning programmes (see also Callender, 1992; CAVTL, 2014). The same principle would apply when standards are translated into certification arrangements, which suggests that standards ought to be developed in collaboration with awarding organisations too. Stanton returned to this theme a few years later, arguing that one of the reasons why agency-designed qualifications have failed in the past was:

The use of linear rather than iterative development processes, without piloting, and with teachers being regarded as implementers of schemes rather than having a role in their design.

(Stanton, Morris & Norrington, 2015, page 79)

In terms of supporting key protagonists, Stanton (2012) also emphasised the risks associated with handing standards over to teachers and trainers without further elaboration, as though they somehow constituted a programme of learning, which is how (in the absence of satisfactory support) they were often treated. Teachers and trainers were required to become qualified in NVQ assessment principles and practices, but not in NVQ teaching and learning principles and practices.¹⁷³

Similarly, the speed with which GNVQs were rolled out left too little time for developing effective support systems and guidance materials for curriculum and

¹⁷³ To be fair, even the assessor qualifications appear to have been something of an afterthought (Raggatt, 1991).

pedagogy (FEU, 1994). Nor was there adequate time or resource for staff development. Indeed, although materials became available over time, including materials designed to facilitate the consistent interpretation of standards, Higham (2003) found little evidence of any systematic form of staff development, other than teachers having to complete units on assessment. Once again, these were units on assessment, not teaching and learning, despite mastery certification being premised on mastery learning, which implicated an entirely different philosophy of teaching and learning, not to mention entirely different teaching and learning strategies (see also FEFC, 1997). As one of the awarding organisation officials interviewed by Kathryn Ecclestone put it:

So you had school teachers blundering into this with a lack of specialist knowledge, with no materials and a style of learning that they had never had to use before.

(Ecclestone, 2000, page 549)

Even toward the end of the 1990s, Ofsted was still insisting that GNVQ teachers needed more help with curriculum planning, teaching methodology, and assessment (Ofsted, 1999). Butcher (1998) described how Initial Teacher Training paid scant attention to the particular requirements of GNVQ. More generally, Ecclestone (2010) emphasised how the state of Continuing Professional Development continued to be extremely poor. These challenges were compounded by the more fundamental challenge of appointing suitably qualified teachers and trainers in the further education sector (Wheatley, 1959; FEFC, 1997; Lingfield, 2012).

It was not just the NVQ and GNVQ reforms that failed to involve and support teachers and trainers adequately from the outset. The same was true when outcome-based qualifications were first introduced by the TEC and the BEC. Again, these were radically new qualifications, which changed what it meant to teach and learn, as well as carrying an expectation of needing to be customised to local needs. None of this was adequately factored into the qualification reform process. Former Minister of State for Education and Science, Gerry Fowler, put it like this:

This takes me back to BEC and TEC. In my experience, such is the burden upon some colleges of devising new syllabuses for the approval of the Councils that there will have to be either an acceptance of the need to use (in the short term at least) prepared syllabuses and even prepackaged course materials, or a recognition by the DES and Local Education Authorities that bricks cannot be built without straw. Additional staff resources would solve the problem; their provision seems unlikely. There is a lesson in this for the future. Further changes in the machinery of course validation may be highly desirable, as I have argued. But they will have to come gradually and piecemeal, if the system is not to collapse under the strain.

(Fowler, 1978, page 57)

Some years after the introduction of the new BTEC model, an evaluation conducted by the National Foundation for Educational Research emphasised that many staff still lacked necessary skills. The evaluation report explained that there was a “need for radical development of the complex of skills required for the role of manager and facilitator of learning” (FEU & BTEC, 1990, page 7). An earlier report from the Further Education Unit had raised similar concerns (FEU, 1986).¹⁷⁴

This concluding section can be summarised in a quotation from Alan Brown, who explained that the clear lesson from framework-driven reforms in England is that:

an emphasis on qualifications development needs to be balanced with equal concern about how learning and development will be facilitated in practice

(Brown, 2011, page 04-2)

Brown is just one of many scholars who have reached a similar conclusion (see Hodkinson & Issitt, 1995, for another example). It remains unclear why this lesson was not straightforwardly learnt from earlier TVET reforms to later ones, although a lack of policy memory may have something to do with this.

Conclusion

It is hard to draw a pithy conclusion from our nuanced and multifaceted investigation into the origins and evolution of the CASLO approach in England. So, we will end with just a couple of high-level observations concerning the CASLO approach in the context of outcome-based and mastery-based qualifications more generally.

First, the very fact that we can distinguish the CASLO approach so clearly – as a high-level design template that has underpinned so many TVET qualifications over the past few decades – is interesting and important in its own right. As we have seen, the approach is just one way of operationalising an outcome-based qualification model, and just one way of operationalising a mastery-based qualification model. So, why did it become ‘the’ way in England, achieving almost hegemonic status as a TVET qualification model?

Clearly, the answer has much to do with the homogenising tendency of qualification regulation in England, which led to the approach being specified as accreditation criteria within both NVQ and QCF regulations. However, it also has something to do with the general zeitgeist of enthusiasm for outcome-based and mastery-based

¹⁷⁴ Of course, in the wake of the withdrawal of the QCF, where qualifications began to transition away from the CASLO approach, essentially the same challenge arose with the need to support teachers and trainers in delivering more classically designed qualifications.

approaches, which took root during the 1970s in the wake of the Objectives Movement, and which persisted within the TVET qualification sector despite not taking root in the general qualification sector nor in the national curriculum assessment context. Once the CASLO approach rose to national prominence within the NVQ model, it set a precedent for qualification designers to follow. Some followed under duress. Others followed willingly.

Second, given its roots in the Mastery Movement, the CASLO approach is fundamentally a philosophy of teaching, learning, and assessment, as embodied in certification requirements. In other words, it is not merely the 'occupational way' of designing qualifications. Of course, it makes a lot of sense in certain occupational contexts, where achieving mastery means becoming fully competent in an occupational role. Yet, even in occupational contexts, it is still premised upon mastery learning, so the underpinning philosophy is still there, albeit only implicitly.

Contrary to conclusions from certain policy reviews, the CASLO approach is not straightforwardly inappropriate in general education contexts. Far from it. The very idea of mastery learning was a response to the tendency, within general education, for teachers to treat failure as normal and unproblematic. However, it is entirely legitimate to question whether the CASLO approach is appropriate for any particular general education context. Indeed, it might well be judged appropriate for certain contexts but not for others. Furthermore, it is quite possible to adopt a mastery approach to teaching and learning without that necessarily culminating in mastery certification. Finally, the CASLO approach incorporates a plausible approach to mastery certification, but it is not the only approach. So, this begs the question of whether it is the optimal approach for any particular context.

Over the years, we have come to conceptualise and operationalise outcome-based qualifications quite narrowly. Hence, the CASLO approach achieving almost hegemonic status. Yet, there is nothing sacrosanct about it. There are certainly better and worse ways of implementing the CASLO approach, but there might also be better ways of designing outcome-based qualifications, more generally, and better ways of certifying mastery. This behoves us to think more broadly and creatively about the significance of outcomes and mastery when designing vocational and technical qualifications for the future.

References

- Abbott, M.K., McLone, R.R., & Patrick, H. (1989). *GCSE Inter-Group Comparability Study 1988: English*. Organised by the Midland Examining Group on behalf of the Inter-Group Research Committee for the GCSE and the Joint Council for the GCSE.
- Acland, A.H.D. (1911). *Report of the Consultative Committee on Examinations in Secondary Schools*. Cd 6004. London: HMSO.
- Ainley, P. (1990). *Vocational Education and Training*. London: Cassell Educational Limited.
- Allen, M. (2004). *The Rise and Fall of the GNVQ: A study of the changing relationship between young people and vocational qualifications at the start of the twenty-first century*. PhD thesis. Milton Keynes: The Open University.
- Anderson, K. (1984). Implications of BEC philosophy. *Journal of Further and Higher Education*, 8(2), 14–22.
- Anderson, L.W., Krathwohl, D.R., Airasian, P.W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., Raths, J., & Wittrock, M.C. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A revision of Bloom's Taxonomy of Educational Objectives* (Complete edition). New York: Longman.
- Anson, R.W. (1978). National qualifications in cartography: A review. *The Cartographic Journal*, 15(2), 66–71.
- Arad Research Ltd. (2014). *The Credit and Qualifications Framework for Wales: A qualitative review of its impact*. 73/2014. Cardiff: Welsh Government.
- Ashworth, P. (1992). Being competent and having 'competencies'. *Journal of Further and Higher Education*, 16(3), 8–17.
- Ashworth, P.D. & Saxton, J. (1990). On 'Competence'. *Journal of Further and Higher Education*, 14(2), 3–25.
- Bacon, F. (1969). Continuous assessment. *Technical Education (Education + Training)*, 11(7), 282–283.
- Bates, I. (1995). The competence movement and the national vocational qualification framework: The widening parameters of research. *British Journal of Education and Work*, 8(2), 5–13.
- Beaumont, G. (1996). *Review of 100 NVQs and SVQs*. London: National Council for Vocational Qualifications.
- Bell, J. (1968). Course and examinations for technicians. *Technical Education*, 10(8), 327–329.

- Berk, R.A. (1978) (Ed.). *Criterion-Referenced Measurement: The state of the art*. Baltimore, MD: John Hopkins University Press.
- Birbeck, R. (1980). The history and operation of the Technician Education Council programmes in colleges of technology in the United Kingdom. *International Journal of Electrical Engineering Education*, 17(4), 293–299.
- BIS & DfE (Department for Business, Innovation and Skills & Department for Education) (2016). *Post-16 Skills Plan*. Cm 9280. London: HMSO.
- BIS (Department for Business, Innovation and Skills) (2012). *New Challenges, New Chances. Further Education and Skills System Reform Plan: Building a world class skills system*. London: Department for Business, Innovation and Skills.
- BIS (Department for Business, Innovation and Skills) (2014). *Getting the Job Done: The government's reform plan for vocational qualifications*. London: Department for Business, Innovation and Skills.
- Black, H. & Wolf, A. (1990). *Knowledge and Competence: Current issues in training and education*. London: Careers and Occupational Information Centre.
- Blakey, H. & Stagg, B. (1978). The Technician Education Council. *The Post Office Electrical Engineers' Journal*, 70(4), 219–226.
- Block, J.H. (1971). Introduction to mastery learning: Theory and practice. In J.H. Block (Ed.). *Mastery Learning: Theory and Practice* (pp.2–12). New York: Holt, Rinehart, and Winston Inc.
- Bloom, B.S. (1968). Learning for mastery. *Evaluation Comment*, 1(2), 1–12.
- Bloom, B.S., Englehart, M.D., Furst, E.J., Hill, W.H., & Krathwohl, D.R. (1956). *Taxonomy of Educational Objectives: The classification of educational goals. Handbook I: Cognitive domain*. London: Longmans, Green and Co Ltd.
- Bloom, B.S., Hastings, J.T., & Madaus, G.F. (1971). *Handbook on Formative and Summative Evaluation of Student Learning*. New York: McGraw-Hill Book Company.
- Bobbitt, F. (1918). *The Curriculum*. Boston: Houghton Mifflin.
- Bobbitt, F. (1924). *How to Make a Curriculum*. Boston: Houghton Mifflin.
- Boffy, R. (1990). Occupational competence and work-based learning: The future for further education. In M. Bees & M. Swords (Eds.). *National Vocational Qualifications and Further Education* (pp.182–199). London: Kogan Page Ltd.
- Bolton, W. (1978). Forwards with TEC. *Journal of Further and Higher Education*, 2(2), 29–35.
- Bourne, R. (1984). The rise and fall of the HNC: Long live the HNC! *IEE Proceedings A*, 131(9), 745–748.

- Bowe, R. & Whitty, G. (1984). Teachers, boards and standards: The attack on school-based assessment in English public examinations at 16+. In P. Broadfoot (Ed.). *Selection, Certification & Control: Social issues in educational assessment* (pp.179–197). Sussex: The Falmer Press.
- Bracknell/Wokingham School-Industry Partnership (1980). *A Simple Guide to BEC and TEC*. Slough: South Bucks & East Berks Chamber of Commerce & Industry.
- Bridgman, P.W. (1927). *The Logic of Modern Physics*. New York: Macmillan.
- Briggs, D.C. (2022). *Historical and Conceptual Foundations of Measurement in the Human Sciences: Credos and controversies*. New York: Routledge.
- Broadfoot, P. (1986) (Ed.). *Profiles and Records of Achievement: A review of issues and practice*. East Sussex: Holt, Rinehart and Winston Ltd.
- Brockmann, M., Clarke, L., & Winch, C. (2008). Can performance-related learning outcomes have standards? *Journal of European Industrial Training*, 32(2/3), 99–113.
- Brockmann, M., Clarke, L., & Winch, C. (2008). Competence and competency in the EQF and in European VET systems. *Journal of European Industrial Training*, 33(8/9), 787–799.
- Brown, A. (2011). Problems with National Qualifications Frameworks in Practice: The English case. *Magazin erwachsenenbildung.at*, 14, 04-1–04-11.
- Brown, M. (1994). An introduction to the discourse on competency-based training (CBT). In M. Brown, F. Patrick, D. Tate, & S. Wright (1994) (Eds.). *A Collection of Readings Related to Competency-Based Training* (pp.1–17). Australia: Deakin University Faculty of Education.
- BTEC (Business and Technician Education Council) (1984). *Policies and Priorities into the 1990s*. London: Business and Technician Education Council.
- BTEC (Business and Technician Education Council) (1985). *BTEC Bulletin No. 3* (April). London: Business and Technician Education Council.
- BTEC (Business and Technician Education Council) (1986). *Assessment and Grading: General guideline*. London: Business and Technician Education Council.
- BTEC (Business and Technician Education Council) (1989). *BTEC Bulletin No. 14* (January). London: Business and Technician Education Council.
- Bull, H. (1985). The use of behavioural objectives: A moral issue? *Journal of Further and Higher Education*, 9(3), 74–80.
- Burke, J. (1989) (Ed.). *Competency Based Education and Training*. Sussex: The Falmer Press.

- Burke, J. (1995) (Ed.). *Outcomes, Learning and the Curriculum: Implications for NVQ's, GNVQ's and other qualifications*. London: The Falmer Press.
- Butcher, J. (1998) The conundrum of General National Vocational Qualifications: A case study of the training needs of school-based GNVQ teachers. *Journal of Vocational Education and Training*, 50(4), 569–583.
- Butterfield, S. (1996). *Educational Objectives and National Assessment*. Buckingham: Open University Press.
- Callender, C. (1992). *Will National Vocational Qualifications Work? Evidence from the construction industry*. IMS Report No 228. East Sussex: Institute of Manpower Studies.
- Cantor, L., Roberts, I., & Pratley, B. (1995). *A Guide to Further Education in England and Wales*. London: Cassell.
- Cantor, L.M. & Roberts, I.F. (1972). *Further Education in England and Wales* (2nd edition). London: Routledge & Kegan Paul Ltd.
- Capey, J. (1995). *GNVQ Assessment Review: Final report of the review group chaired by Dr John Capey*. London: National Council for Vocational Qualifications.
- Carr, R. (1958). *Training for Skill: Recruitment and training of young workers in industry. Report by a sub-committee of the National Joint Advisory Council*. London: HMSO.
- Carroll, G. & Boutall, T. (2011). *Guide to Developing National Occupational Standards: Revised June 2011*. London: UK Commission for Employment and Skills.
- Carroll, J.B. (1963). A model of school learning. *Teachers College Record*, 64(8), 723–733.
- Carter, A. (2012). *Assessment-in-Action: A study of lecturers' and students' constructions of BTEC national assessment practice in a college engineering programme area*. EdD thesis. Bristol: School of Education, University of the West of England.
- Carter, A. & Bathmaker, A.M. (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.
- Cassels, J. (2001). *Modern Apprenticeships: The way to work. A report of the Modern Apprenticeship Advisory Committee*. London: Department for Education and Skills.

- CAVTL (Commission on Adult Vocational Teaching and Learning) (2013). *It's About Work... Excellent adult vocational teaching and learning*. London: Learning and Skills Improvement Service.
- Cedefop (2009). *The Shift to Learning Outcomes: Policies and practices in Europe*. Luxembourg: Office for Official Publications of the European Communities.
- Cedefop (2016). *Application of Learning Outcomes Approaches Across Europe: A comparative study*. Luxembourg: Publications Office of the European Union.
- Cedefop (2017). *Defining, writing and applying learning outcomes: A European handbook*. Luxembourg: Publications Office of the European Union.
- Christie, R. (1985). Training for competence: The nature and assessment of the beast. *Journal of European Industrial Training*, 9(6), 30–32.
- City & Guilds (2014). *Sense and Instability: Three decades of skills and employment policy*. London: City & Guilds Group.
- City & Guilds (2016). *Sense and Instability 2016*. London: City & Guilds Group.
- City & Guilds (2019). *Sense and Instability 2019*. London: City & Guilds Group.
- Clegg, C.R. & Jones, A. (1970). Skills Testing Service. *Industrial and Commercial Training*, 2(6), 252–255.
- Cook, A. (2001). Testing the water: Practitioner opinion of a regional credit scheme (NICATS). *Higher Education Quarterly*, 55(3), 239–256.
- Crellin, J.R., Orton, R.J.J., & Tawney, D.A. (1979). Present-day school physics syllabuses. *Reports on Progress in Physics*, 42(4), 677–725.
- Cresswell, M.J. (1987). Describing examination performance: Grade criteria in public examinations. *Educational Studies*, 13(3), 247–265.
- Crowther, G. (1959). *15 to 18: A report of the Central Advisory Council for Education (England). Volume I*. London: HMSO.
- 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.
- Daugherty, R. (1995). *National Curriculum Assessment: A review of policy 1987–1994*. London: The Falmer Press.
- Davies, A. (1981) BEC General Awards: An approach to national standards. *Journal of Further and Higher Education*, 5(2), 57–68.
- Davies, I.K. (1976). A history of the objectives movement in education. In M. Brown, F. Patrick, D. Tate, & S. Wright (1994) (Eds.). *A Collection of Readings Related to Competency-Based Training* (pp.75–90). Australia: Deakin University Faculty of Education.

- De Ville (1986). *Review of Vocational Qualifications in England and Wales*. London: HMSO.
- Dearing, R. (1993). *The National Curriculum and its Assessment: Final report*. London: School Curriculum and Assessment Authority.
- Dearing, R. (1996). *Review of Qualifications for 16 to 19 Year Olds: Full report*. London: School Curriculum and Assessment Authority.
- Debling, G. & Hallmark, A. (1990). Identification and assessment of underpinning knowledge and understanding in the context of the UK government's standards programme. In H. Black & A. Wolf (Eds.). *Knowledge and Competence: Current issues in training and education* (pp.8–11). London: Careers and Occupational Information Centre.
- Debling, G. (1989). The Employment Department/Training Agency Standards Programme and NVQs: Implications for education. In J. Burke (Ed.). *Competency Based Education and Training* (pp.77–94). Sussex: The Falmer Press.
- Debling, G. (1990). Foreword. In H. Black & A. Wolf (Eds.). *Knowledge and Competence: Current issues in training and education* (pp.2–2). London: Careers and Occupational Information Centre.
- Debling, G. (1991). Developing standards. In P. Raggatt & L. Unwin (Eds.). *Change and Intervention: Vocational education and training* (pp.1–21). London: The Falmer Press.
- DES & DE (Department of Education and Science & Department of Employment) (1985). *Education and Training for Young People*. Cmnd 9482. London: HMSO.
- DES & WO (Department of Education and Science & Welsh Office) (1982). *Examinations at 16-Plus: A statement of policy*. London: HMSO.
- DES (Department of Education and Science) (1985). *GCSE National Criteria*. London: HMSO.
- DES, DE, & WO (Department of Education and Science, Department of Employment, & Welsh Office) (1991). *Education and Training for the 21st Century*. Cm 1536 Volume I. London: HMSO.
- DfE & BIS (Department for Education & Department for Business, Innovation and Skills) (2013). *The Future of Apprenticeships in England: Next steps from the Richard review*. London: Department for Business, Innovation and Skills.
- DfE (Department for Education) (2011a). *Wolf Review of Vocational Education: Government response*. London: Department for Education.

- DfE (Department for Education) (2011b). *Qualifications for 14 – 16 Year Olds and Performance Tables: Policy paper for consultation*. London: Department for Education.
- DfE (Department for Education) (2011c). *Qualifications for 14 – 16 Year Olds and Performance tables: Technical guidance for awarding organisations*. London: Department for Education.
- DfE (Department for Education) (2013a). *Government Proposals to Reform Vocational Qualifications for 16-19 Year Olds: Launch date 7 March 2013*. London: Department for Education.
- DfE (Department for Education) (2013b). *16 19 Accountability Consultation: Launch date 12 September 2013*. London: Department for Education.
- DfE (Department for Education) (2017). *Technical and Applied Qualifications for 14 to 19 Year Olds. Key stage 4 and 16 to 18 performance tables from 2020: technical guidance for awarding organisations*. London: Department for Education.
- DfE (Department for Education) (2019). *Review of Post-16 Qualifications at Level 3 and Below in England: Government consultation*. London: Department for Education.
- DfEE (Department for Education and Employment) (2000). *Opportunity for All: Skills for the new economy. Initial response to the National Skills Task Force final report from the Secretary of State for Education and Employment*. London: Department for Education and Employment.
- DfES (Department for Education and Skills) (2001). *Meeting the Sector Skills and Productivity Challenge: Building a stronger sector skills network*. London: Department for Education and Skills.
- DfES (Department for Education and Skills) (2002). *Success for All: Reforming Further Education and training. Discussion document: June 2002*. London: Department for Education and Skills.
- DfES (Department for Education and Skills) (2003a). *21st Century Skills: Realising our potential. Individuals, employers, nation*. Cm 5810. London: The Stationery Office.
- DfES (Department for Education and Skills) (2003b). *Government Response to the Committee's Third Report: A Level Standards. House of Commons Education and Skills Committee. Third Special Report of Session 2002–03*. HC 1026. London: The Stationery Office Limited.
- DIUS & DCSF (Department for Innovation, Universities and Skills & Department for Children, Schools and Families) (2008). *World-class Apprenticeships: Unlocking Talent, Building Skills for All. The Government's strategy for the*

- future of Apprenticeships in England*. London: Department for Innovation, Universities and Skills.
- Ecclestone, K. (1992). *Understanding Accreditation: Ways of recognising achievement*. London: Further Education Unit.
- Ecclestone, K. (1997) On the frontline or at the margins? Finding a place in post-compulsory research. *Research in Post-Compulsory Education*, 2(3), 299–312.
- Ecclestone, K. (1998) Euston Road and the ivory towers: The impact of the GNVQ research industry. *Journal of Education Policy*, 13(6), 679–697.
- Ecclestone, K. (2000) Bewitched, bothered and bewildered: A policy analysis of the GNVQ assessment regime 1992–2000. *Journal of Education Policy*, 15(5), 539–558.
- Ecclestone, K. (2002). *Learning Autonomy in Post-16 Education: The politics and practice of formative assessment*. Oxford: RoutledgeFalmer.
- Ecclestone, K., with Davies, J., Derrick, J. & Gawn, J. (2010). *Transformative Formative Assessment in Lifelong Learning*. Berkshire: Open University Press.
- ED (Employment Department) (1992). *Competence & Assessment: Compendium No.2*. London: Employment Department.
- ED (Employment Department) (1994). *The Place of Knowledge and Understanding in the Development of National Vocational Qualifications and Scottish Vocational Qualifications*. London: Employment Department.
- Edmonds, T. & Stuart, D. (1992). Assessing competence at higher levels: The management experience. In Employment Department (Ed.). *Competence & Assessment: Compendium No.2* (pp.49–51). London: Employment Department.
- Elliott, G. (2001). An alternative model of competence-based assessment. In C. Velde (Ed.). *International Perspectives on Competence in the Workplace* (pp.85–98). Dordrecht, Netherlands: Springer.
- Elliott, J. (1979). BEC is on course. *Education & Training (Education + Training)*, 21(8), 227–228.
- Ellis, P. (1979). Achievement testing. *Journal of European Industrial Training*, 3(7), 17–21.
- Ellison, J.R. (1987). *An Integrated Studies Approach to Business Education with Special Reference to the Theory and Practice of Business Education Council Courses*. MA Thesis. Durham: University of Durham.

- Eraut, M., Steadman, S., Trill, J., & Parkes, J. (1996). *The Assessment of NVQs*. Brighton: University of Sussex Institute of Education.
- Evans, R. (2009). [A Short history of Technical Education. Chapter 13: Developments in the 1960s and the 1970s.](#)
- FAB & JCQ (Federation of Awarding Bodies & Joint Council for Qualifications) (2010). *Writing QCF Units: How much detail to provide*. London: Federation of Awarding Bodies.
- FEFC & Ofsted (Further Education Funding Council & Office for Standards in Education) (1999). *GNVQs: Evaluation of the pilot of the new assessment model, 1997 to 1999*. Coventry: Further Education Funding Council.
- FEFC (Further Education Funding Council) (1994a). *National Vocational Qualifications in the Further Education Sector in England: National survey report*. Coventry: Further Education Funding Council.
- FEFC (Further Education Funding Council) (1994b). *General National Vocational Qualifications in the Further Education Sector in England: National survey report*. Coventry: Further Education Funding Council.
- FEFC (Further Education Funding Council) (1995). *General National Vocational Qualifications in the Further Education Sector in England*. Coventry: Further Education Funding Council.
- FEFC (Further Education Funding Council) (1997). *Standards and their Assurance in Vocational Qualifications*. Coventry: Further Education Funding Council.
- Fennell, E. (1991). *Development of Assessable Standards for National Certification*. London: Employment Department.
- FEU & BTEC (Further Education Unit & Business and Technician Education Council) (1990). *Teaching and Learning Strategies: An evaluation of BTEC courses*. London: Further Education Unit.
- FEU (Further Education Unit) (1979). *A Basis for Choice: Report of a study group on post-16 pre-employment courses*. London: Department of Education and Science.
- FEU (Further Education Unit) (1984). *Towards a Competence-Based System: An FEU view*. London: Further Education Unit.
- FEU (Further Education Unit) (1986). *Assessment, Quality and Competence: An FEU view*. London: Further Education Unit.
- FEU (Further Education Unit) (1992). *A Basis for Credit?* London: Further Education Unit.

- FEU (Further Education Unit) (1993). *A Basis for Credit? Developing a Post-16 Credit Accumulation and Transfer Framework: Feedback and developments*. London: Further Education Unit.
- FEU (Further Education Unit) (1994). *Implementing GNVQs: A manual*. London: Further Education Unit.
- FEU (Further Education Unit) (1995a). *A Framework for Credit: A common framework for post-14 education and training for the twenty-first century*. London: Further Education Unit.
- FEU (Further Education Unit) (1995b). *Framework for Credit. Framework guidelines 2. Learning outcomes, units and modules*. London: Further Education Development Agency.
- Field, J. (1991). Competency and the pedagogy of labour. In M. Brown, F. Patrick, D. Tate, & S. Wright (1994) (Eds.). *A Collection of Readings Related to Competency-Based Training* (pp.283–292). Australia: Deakin University Faculty of Education.
- Field, J. (1995). Reality testing in the workplace: Are NVQs ‘Employment-led’? In P. Hodgkinson & M. Issitt (Eds.). *The Challenge of Competence: Professionalism through vocational education and training* (pp.28–43). London: Cassell.
- Field, S. (2018). *The Missing Middle: Higher Technical Education in England*. London: Gatsby Charitable Foundation.
- Fishbein, J.M. & Tyler, R. (1973). The father of behavioral objectives criticizes them: An interview with Ralph Tyler. *The Phi Delta Kappan*, 55(1), 55–57.
- Fisher, R. (1999). *The Vocational Curriculum in England 1974–1994: A socio-historical study of the Business and Technology Education Council's National Diploma in Business and Finance*. PhD thesis. Huddersfield: University of Huddersfield.
- Fisher, R. (2003). The golden age of BTEC: the business education curriculum in 1980s further education. *The Curriculum Journal*, 14(2), 253–277.
- Fisher, R. (2004). From Business Education Council to Edexcel Foundation 1969–96: The short but winding road from technician education to instrumentalist technicism. *Journal of Education and Work*, 17(2), 237–255.
- FISSS (Federation for Industry Skills & Standards) (2014). *Employer-defined Apprenticeship Standards. A toolkit for enablers and facilitators of Trailblazer groups*. Edinburgh: Federation for Industry Sector Skills & Standards.
- Flew, A. (1979). *A Dictionary of Philosophy*. London: Pan Books Ltd.

- Flyvbjerg, B. & Gardner, D. (2023). *How Big Things Get Done: The surprising factors behind every successful project, from home restoration to space exploration*. London: Macmillan.
- Foden, F.E. (1951). The National Certificate. *The Vocational Aspect of Education*, 3(6), 38–46.
- Foden, F.E. (1966). Technical examinations in England. *Paedagogica Historica*, 6(1), 68–97.
- Fowler, G. (1978). The validation game. *Education & Training (Education + Training)*, 20(2), 56–57.
- Frankland, G. & Ebrahim, N. (2001). *Taking Part One GNVQ Forward*. London: Learning and Skills Development Agency.
- Franklin, R., Rawlings, G., & Craven, B. (1983). Barking up the wrong tree? An assessment of BEC's aims and objectives. *Journal of Further and Higher Education*, 7(1), 54–61.
- Fuller, A. & Unwin, L. (2009). Change and continuity in apprenticeship: The resilience of a model of learning. *Journal of Education and Work*, 22(5), 405–416.
- Furst, E.J. (1981). Bloom's Taxonomy of Educational Objectives for the Cognitive Domain: Philosophical and educational issues. *Review of Educational Research*, 51(4), 441–453.
- Gagne, R.M. (1988). Mastery learning and instructional design. In R.C. Richey (Ed.). *The Legacy of Robert M. Gagne* (pp.107–126). Washington, DC: US Department of Education.
- Gallacher, J., Toman, N., Caldwell, J., Raffe, D., & Edwards, R. (2005). *Evaluation Of The Impact Of The Scottish Credit And Qualifications Framework (SCQF)*. Glasgow: Department of Enterprise, Transport and Lifelong Learning of the Scottish Executive.
- GCE Boards (1983). *Common Cores at Advanced Level*. Various locations: The GCE Boards of England, Wales and Northern Ireland.
- Gipps, C. (1990). *Assessment: A teachers' guide to the issues*. London: Hodder & Stoughton.
- Glaser, R. (1963). Instructional technology and the measurement of learning outcomes. *American Psychologist*, 18(8), 519–521.
- Gokulsing, K., Ainley, P., & Tysome, T. (1996). *Beyond Competence: The National Council for Vocational Qualifications framework and the challenge to Higher Education in the new millennium*. Hampshire: Avebury Ashgate Publishing Limited.

- Greatorex, J. (2001). *Can vocational A levels be meaningfully compared with other qualifications?* Paper presented at the Annual Conference of the British Educational Research Association Conference, 13–15 September, University of Leeds.
- Gronlund, N.E. & Brookhart, S.M. (2008). *Gronlund's Writing Instructional Objectives* (8th edition). Upper Saddle River, New Jersey: Pearson Education.
- Grugulis, I. (2002). *Skill and Qualification: The contribution of NVQs to raising skill levels*. SKOPE Research Paper No.36. Coventry: Centre on Skills, Knowledge and Organisational Performance, University of Warwick.
- Guskey, T.R. (2023). *Implementing Mastery Learning* (3rd edition). Thousand Oaks, CA: Corwin.
- Halliday, J. (2004). Competence in the workplace: Rhetorical robbery and curriculum policy. *Educational Philosophy and Theory*, 36(5), 579–590.
- Halliday, J.S. (1981). TEC physics assessment. *Physics Education*, 16(3), 172–177.
- Hannagan, T.J. (1978). What's in a Name? A Reappraisal of the Higher National Diploma in Business Studies. *Journal of Further and Higher Education*, 2(3), 88–92.
- Hansen, G.B. (1967). *Britain's Industrial Training Act: Its history, development and implications for America*. Washington, D.C.: National Manpower Policy Taskforce.
- Hargraves, G. (1995). The influence of the European communities on the emergence of competence-based models of vocational training in England and Wales. *British Journal of Education and Work*, 8(2), 28–40.
- Hargraves, G. (2000). The Review of Vocational Qualifications, 1985 to 1986: An analysis of its role in the development of competence-based vocational qualifications in England and Wales. *British Journal of Educational Studies*, 48(3), 285–308.
- Haslegrave, H.L. (1969). *Report of the Committee on Technician Courses and Examinations*. London: HMSO.
- HCESC (House of Commons Education and Skills Committee) (2003). *A Level Standards: Third report of session 2002–3*. HC 153. London: The Stationery Office.
- HEFCE (Higher Education Funding Council for England) (2015). *Young Participation in Higher Education: A-levels and similar qualifications*. Bristol: Higher Education Funding Council for England.

- Higginson, G. (1988). *Advancing A Levels: Report of a committee appointed by the Secretary of State for Education and Science and the Secretary of State for Wales*. London: HMSO.
- Higham, J. & Yeomans, D. (2007). Policy memory and policy amnesia in 14–19 education: Learning from the past? In D. Raffe & K. Spours (Eds.). *Policy-making and Policy Learning in 14–19 Education* (pp.33–60). London: Institute of Education, University of London.
- Higham, J. (2003). Curriculum change: A study of the implementation of General National Vocational Qualifications. *The Curriculum Journal*, 14(3), 327–350.
- Higham, J., Sharp, P., & Yeomans, D. (1996). *The Emerging 16–19 Curriculum: Policy and provision*. London: David Fulton Publishers Ltd.
- Hirst, P.H. (1969). The Logic of the Curriculum. *Journal Of Curriculum Studies*, 1(2), 142–158.
- Hlebowitsh, P.S. (1992). Amid behavioural and behaviouristic objectives: Reappraising appraisals of the Tyler Rationale. *Journal of Curriculum Studies*, 24(6), 533–547.
- Hlebowitsh, P.S. (1995). Interpretations of the Tyler Rationale: A reply to Kliebard, *Journal of Curriculum Studies*, 27(1), 89–94.
- HM Government (2010). *The Coalition: Our programme for government*. London: Cabinet Office.
- HM Government (2013a). *The Future of Apprenticeships in England: Implementation plan*. London: Department for Business, Innovation and Skills.
- HM Government (2013b). *The Future of Apprenticeships in England: Guidance for Trailblazers. Version 1 – October 2013*. London: Department for Business, Innovation and Skills.
- HM Government (2014). *The Future of Apprenticeships in England: Guidance for Trailblazers. Version 2 – March 2014*. London: Department for Business, Innovation and Skills.
- HM Government (2015). *The Future of Apprenticeships in England: Guidance for Trailblazers – from standards to starts*. London: Department for Business, Innovation and Skills.
- 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.
- Hodgson, A. & Spours, K. (2003). *Beyond A levels: Curriculum 2000 and the reform of 14–19 qualifications*. London: Kogan Page Limited.

- Hodkinson, P. & Issitt, M. (1995) (Eds.). *The Challenge of Competence: Professionalism through vocational education and training*. London: Cassell.
- Hodkinson, P. (1992). Alternative models of competence in vocational education and training. *Journal of Further and Higher Education*, 16(2), 30–39.
- Hooper, R. (1971). Introduction to part II. In R. Hooper (Ed.). *The Curriculum: Context, design & development* (pp.116–123). Edinburgh: Oliver & Boyd, with The Open University Press.
- Huddleston, P. (2020). *A Short History of Employer Engagement: Once more round the buoy or set fair for a better voyage?* London: Education and Employers.
- Humphreys, J. (1981) Towards the in-college evaluation of TEC programmes. *Journal of Further and Higher Education*, 5(1), 40–50.
- Hunter, D.W. (1985). The Teaching of Electronics in Schools and Further Education: A case study in curriculum change. MA thesis. Durham: University of Durham.
- Hyland, T. (1993). Competence, knowledge, and education. *Journal of Philosophy of Education*, 27(1), 57–68.
- Hyland, T. (1994). *Competence, Education and NVQs: Dissenting perspectives*. London: Cassell.
- James, S. (2006). *Adding New Ingredients to an Old Recipe: NVQs and the influence of CATERBASE*. SKOPE Issues Paper 13. Oxford: Centre on Skills, Knowledge and Organisational Performance, University of Oxford.
- Jessup, G. & Jessup, H. (1975). *Selection and Assessment at Work*. London: Methuen & Co Ltd.
- Jessup, G. (1989). The emerging model of vocational education and training. In J. Burke (Ed.). *Competency Based Education and Training* (pp.65–76). Sussex: The Falmer Press.
- Jessup, G. (1990). National Vocational Qualifications: Implications for further education. In M. Bees & M. Swords (Eds.). *National Vocational Qualifications and Further Education* (pp.17–33). London: Kogan Page Ltd.
- Jessup, G. (1991). *Outcomes: NVQs and the emerging model of education and training*. London: The Falmer Press.
- Jessup, G. (1993). Towards a coherent post-16 qualifications framework: The role of GNVQs. In W. Richardson, J. Woolhouse, & D. Finegold (Eds). *The Reform of Post-16 Education and Training in England and Wales* (pp.131–145). Harlow, Essex: Longman Information & Reference.
- Joint Awarding Body Steering Group (2001). *Joint Awarding Body Guidance on Internal Verification of NVQs*. London: Department for Education and Skills.

- JMB (Joint Matriculation Board) (1970). *Examining in Advanced Level Science Subjects of the GCE*. Occasional Publication 30. Manchester: Joint Matriculation Board.
- Jones, A. & Whittaker, P. (1975). *Testing Industrial Skills*. Essex: Gower Press.
- Jones, A. (1971). Performance tests for industrial skills: Some of their techniques, uses and problems. *Industrial and Commercial Training*, 3(10), 481–484.
- Jones, L. & Moore, R. (1995). Appropriating Competence: the competency movement, the New Right and the ‘culture change’ project. *British Journal of Education and Work*, 8(2), 78–92.
- Joseph, K. (1984). Speech by the Rt Hon Sir Keith Joseph, Secretary of State for Education and Science, at the North of England Education Conference, Sheffield, on Friday 6 January 1984. *Oxford Review of Education*, 10(2), 137–146.
- Kelly, A.V. (1982). *The Curriculum: Theory and practice* (2nd edition). Harper & Row.
- Kelly, A.V. (2009). *The Curriculum: Theory and practice* (6th edition). London: SAGE Publications Ltd.
- Kingdon, M. & Stobart, G. (1988). *GCSE Examined*. Sussex: The Falmer Press.
- Kingdon, M. (1991). *The Reform of Advanced Level*. London: Hodder & Stoughton.
- Klausmeier, K. & Goodwin, W. (1966). *Learning and Human Abilities* (2nd edition). London: Harper Row Ltd.
- Kliebard, H.M. (1995). The Tyler Rationale revisited. *Journal of Curriculum Studies*, 27(1), 81–88.
- Krathwohl, D.R. (2002). A Revision of Bloom’s Taxonomy: An overview. *Theory Into Practice*, 41(4), 212–218.
- Laczik, A. & Fettes, T. (2021). *Perspectives on National Occupational Standards (NOS): What do users think?* Glasgow: Skills Development Scotland.
- Laczik, A., Dabbous, D., Patel, J., & Wilson, E. (2023). *Good Policy Making: What can we learn from past policies?* London: Edge Foundation.
- Latham, H. (1886). *On the Action of Examinations Considered as a Means of Selection*. Boston, MA: Willard Small.
- le Roux, A.A. (1983). Barking up the wrong tree? A comment. *Journal of Further and Higher Education*, 7(3), 3–7.
- Leahey, T.H. (1992). *A History of Psychology: Main currents in psychological thought* (3rd edition). Englewood Cliffs, New Jersey: Prentice Hall.

- LSC & QCA (Learning and Skills Council & Qualifications and Curriculum Authority) (2008). *Qualifications and Credit Framework: Final business case*. Unpublished.
- Leese, J. (1966). From examinations to assessment of student performance. *Technical Education (Education + Training)*, 8(12), 548–550.
- Leitch, S. (2006). *Leitch Review of Skills: Prosperity for all in the global economy – world class skills*. London: The Stationery Office.
- Lester, S. (2011). The UK Qualifications and Credit Framework: A critique. *Journal of Vocational Education & Training*, 63(2), 205–216.
- Lester, S. (2015). Competence standards and frameworks: Some lessons from the United Kingdom. *Edukacja ustawiczna Dorosłych (Polish Journal of Continuing Education)*, 3(90), 132–141.
- Like, C.B. (1986). The BTEC HND in physics – five years on. *Physics Bulletin*, 37(5), 196–197.
- Linford, N. (2015). Trailblazer process ‘out of control,’ says government adviser Nigel Whitehead. [FE Week](#). 11 December.
- Lingfield, R. (2012). *Professionalism in Further Education: Final report of the independent review panel*. London: Department for Business, Innovation and Skills.
- Lockee, B., Moore, D.M., & Burton, J. (2004). Foundations of programmed instruction. In D. Jonassen & M. Driscoll (Eds.). *Handbook of Research on Educational Communications and Technology* (2nd edition) (pp.545–569). New York: Routledge.
- Longbottom, B., Clegg, C., Clarke, P., Turner, M., Fletcher, D., & Hunter, R. (1973). Developing a programme of testing for trainee craftsmen in the shipbuilding industry. *Industrial and Commercial Training*, 5(1), 16–24.
- Lum, G. (2015). On the functions of a system of vocational qualifications. *Journal of Education and Work*, 28(2), 187–206.
- Lysons, K. (1982). BEC – an agent for change. *Education + Training*, 24(4), 98–99.
- Mace, G.R. (1980). The Role of economics in the BEC National Curriculum. *Journal of Further and Higher Education*, 4(3), 62–69.
- MacRory, J., Beaumont, T., & Taylor, H.T. (1977). Backwards from TEC. *Journal of Further and Higher Education*, 1(1), 2–7.
- Mager, R.F. (1962). *Preparing Instructional Objectives*. Palo Alto, CA: Fearon Publishers Inc.

- Mager, R.F. (1984). *Preparing Instructional Objectives* (2nd edition). Belmont, CA: David S. Lake Publishers.
- Mager, R.F. (1997). *Preparing Instructional Objectives: A critical tool in the development of effective instruction* (3rd edition). Atlanta, GA: Center for Effective Performance.
- Maguire, M. (1998). Modern apprenticeships and employers. *Journal of Vocational Education & Training*, 50(2), 247–259.
- Mansfield, B. & Mitchell, L. (1991). The identification of underpinning knowledge – using a knowledge flow chart. In E. Fennell (Ed.). *Development of Assessable Standards for National Certification* (pp.30–34). London: Employment Department.
- Mansfield, B. & Mitchell, L. (1996). *Towards a Competent Workforce*. Aldershot, Hampshire: Gower Publishing Ltd.
- Mansfield, B. (1989). Competence and standards. In J. Burke (Ed.). *Competency Based Education and Training* (pp.26–38). Sussex: The Falmer Press.
- Mansfield, B. (1989). Knowledge, evidence and assessment. In H. Black & A. Wolf (Eds.). *Knowledge and Competence: Current issues in training and education* (pp.15–21). London: Careers and Occupational Information Centre.
- Mansfield, B. (1991). Deriving standards of competence. In E. Fennell (Ed.). *Development of Assessable Standards for National Certification* (pp.12–23). London: Employment Department.
- Mansfield, B. (2004). Competence in transition. *Journal of European Industrial Training*, 28(2/3/4), 296–309.
- Marshall, K. (1991). NVQs: An assessment of the ‘outcomes’ approach to education and training. *Journal of Further and Higher Education*, 15(3), 56–64.
- McCulloch, G. (2001). The reinvention of teacher professionalism. In R. Phillips & J. Furlong (Eds.). *Education, Reform and the State: Twenty-five years of politics, policy and practice* (pp.103–117). London: RoutledgeFalmer.
- Meech, R., with McBirnie, A. & Jones, G. (2018). *Graded Exams in Music, Dance & Drama and the Educational Regulatory Landscape in the UK and Internationally*. London: Council for Dance, Drama and Musical Theatre.
- Melton, R. (1997). *Objectives, Competences & Learning Outcomes: Developing instructional materials in open and distance education*. London: Kogan Page.
- Milloy, A. & Saker, J. (1984). A new approach to BEC assessment. *Journal of Further and Higher Education*, 8(2), 23–27.
- Ministry of Education (1961). *Better Opportunities in Technical Education*. Cmnd 1254. London: HMSO.

- Ministry of Labour (1962). *Industrial Training: Government Proposals*. Cmd 1892. London: HMSO.
- Mirza-Davies, J. (2015). *Apprenticeships Policy, England prior to 2010*. Briefing paper number 07266. London: House of Commons Library.
- Mitchell, L. & Wolf, A. (1991). Understanding the place of knowledge in a competence based approach. In E. Fennell (Ed.). *Development of Assessable Standards for National Certification* (pp.25–29). London: Employment Department.
- Mitchell, L. (1989). The definition of standards and their assessment. In J. Burke (Ed.). *Competency Based Education and Training* (pp.54–64). Sussex: The Falmer Press.
- Montgomery, R.J. (1965). *Examinations: An account of their evolution as administrative devices in England*. London: Longmans, Green & Co Ltd.
- Morris, P. (1977) The Proposals of the Business Education Council: A critical appraisal of BEC as an exercise in curriculum development. *Journal of Further and Higher Education*, 1(3), 3–13.
- Morrison, A.T. (1966). The first cycle – G1 to O2. *The Vocational Aspect of Education*, 18(40), 95–114.
- MSC (Manpower Services Commission) (1981a). *A New Training Initiative: A consultative document*. London: Manpower Services Commission.
- MSC (Manpower Services Commission) (1981b). *A New Training Initiative: An agenda for action*. London: Manpower Services Commission.
- Murtonen, M., Gruber, H., & Lehtinen, E. (2017). The return of behaviourist epistemology: A review of learning outcomes studies. *Educational Research Review*, 22, 114–128.
- Nash, C. (1995). Flexible learning and outcomes. In J. Burke (Ed.). *Outcomes, Learning and the Curriculum: Implications for NVQ's, GNVQ's and other qualifications* (pp.155–168). London: The Falmer Press.
- NCVQ (National Council for Vocational Qualifications) (1987). *The National Vocational Qualification Framework*. London: National Council for Vocational Qualifications.
- NCVQ (National Council for Vocational Qualifications) (1988). *The NVQ Criteria and Related Guidance*. London: National Council for Vocational Qualifications.
- NCVQ (National Council for Vocational Qualifications) (1989). *National Vocational Qualifications: Criteria and procedures*. London: National Council for Vocational Qualifications.

- NCVQ (National Council for Vocational Qualifications) (1991). *Criteria for National Vocational Qualifications*. London: National Council for Vocational Qualifications.
- NCVQ (National Council for Vocational Qualifications) (1993). *The Awarding Bodies Common Accord*. London: National Council for Vocational Qualifications.
- NCVQ (National Council for Vocational Qualifications) (1994). *Grading Advanced GNVQs: Revised criteria and guidance*. London: National Council for Vocational Qualifications.
- NCVQ (National Council for Vocational Qualifications) (1995). *NVQ Criteria and Guidance*. London: National Council for Vocational Qualifications.
- NCVQ (National Council for Vocational Qualifications) (1997a). *The Assessment of Occupational Competence and the Role of Independent Assessment in NVQs/SVQs*. London: National Council for Vocational Qualifications.
- NCVQ (National Council for Vocational Qualifications) (1997b). *External Verification of NVQs: A guide for external verifiers*. London: National Council for Vocational Qualifications.
- NCVQ (National Council for Vocational Qualifications) (1997c). *Assessment of NVQs and SVQs*. London: National Council for Vocational Qualifications.
- Neumann, W. (1979). Educational responses to the concern for proficiency. In M. Brown, F. Patrick, D. Tate, & S. Wright (1994) (Eds.). *A Collection of Readings Related to Competency-Based Training* (pp.91–108). Australia: Deakin University Faculty of Education.
- Newton, P.E. & Lockyer, C. (2022). *How 'CASLO' Qualifications Work*. Ofqual/22/6895. Coventry: Office of Qualifications and Examinations Regulation.
- Newton, P.E. & Shaw, S.D. (2014). *Validity in Educational and Psychological Assessment*. London: SAGE.
- Newton, P.E. (2017). There is more to educational measurement than measuring: The importance of embracing purpose pluralism. *Educational Measurement: Issues and Practice*, 36(2), 5–15.
- Newton, P.E. (2018). *Grading Vocational & Technical Qualifications: Recent policies and current practices*. Ofqual/18/6441/3. Coventry: Office of Qualifications and Examinations Regulation.
- Newton, P.E. (2022). Demythologising A level exam standards. *Research Papers in Education*, 37(6), 875–906.
- Newton, P.E. (2023). Taking purpose into account when designing, using and evaluating educational assessments. In R.J. Tierney, F. Rizvi, & K. Ercikan

- (Eds.). *International Encyclopedia of Education, 4th edition, volume 13* (pp.136–146). Elsevier.
- Nodine, T.R. (2016). How did we get here? A brief history of competency-based higher education in the United States. *Competency-based Education*, 1(1), 5–11.
- Norris, E. & Adam, R. (2017). *All Change: Why Britain is so prone to policy reinvention, and what can be done about it*. London: Institute for Government.
- Norris, N. (1991). The trouble with competence. *Cambridge Journal of Education*, 21(3), 331–341.
- Norwood, C. (1943). *Curriculum and Examinations in Secondary Schools*. London: HMSO.
- NSTF (National Skills Task Force) (2000). *Skills For All: Proposals for a National Skills Agenda. Final Report of the National Skills Task Force*. London: Department for Education and Employment.
- O’Sullivan, G. (1987). Mathematics and statistics on BTEC business courses. *Journal of Further and Higher Education*, 11(3), 130–133.
- Oates, T. (2004). The role of outcomes-based national qualifications in the development of an effective vocational education and training system: the case of England and Wales. *Policy Futures in Education*, 2(1), 53–71.
- Oates, T. (2010). Parity of esteem between academic and vocational qualifications: Time to abandon a misguided notion. In Gatsby Charitable Foundation & Edge Foundation (Ed.). *Technical Education for the 21st Century: Conference report* (pp.14–19). London: Gatsby Charitable Foundation and the Edge Foundation.
- Oates, T. (2013a). *Education, Assessment and Developing a Comprehensive Qualifications Framework*. Cambridge: Cambridge Assessment.
- Oates, T. (2013b). *Towards a New VET: Effective vocational education and training*. Cambridge: Cambridge Assessment.
- Ofqual (2008a). *Regulatory Arrangements for the Qualifications and Credit Framework*. Ofqual/08/3726. London: Office of the Qualifications and Examinations Regulator.
- Ofqual (2008b). *Operating Rules for Using the Term ‘NVQ’ in a QCF Qualification Title*. Ofqual/08/3160. London: Office of the Qualifications and Examinations Regulator.
- Ofqual (2008c). *Arrangements for Awarding and Setting Standards in the Diploma. Principal learning, project and Diploma. Final report of the Diploma Awarding*

and Standards Group. Ofqual/08/3865. London: Office of the Qualifications and Examinations Regulator.

Ofqual (2009a). *Regulating for Confidence in Standards. Consultation on the regulation of qualifications, examinations and assessments*. Ofqual/09/4676. London: Office of the Qualifications and Examinations Regulator.

Ofqual (2009b). *Overview of the 2009 Review and Evaluation of the Qualifications and Credit Framework (QCF)*. London: Office of the Qualifications and Examinations Regulator.

Ofqual (2010a). *Edexcel Level 3 BTEC National Certificate in Manufacturing Engineering: Findings and actions from the monitoring of the qualification*. Ofqual/10/4758. Coventry: Office of Qualifications and Examinations Regulation.

Ofqual (2010b). *The 2009 Evaluation of Qualifications and Credit Framework (QCF) regulation*. Ofqual/10/4644. Coventry: Office of the Qualifications and Examinations Regulator.

Ofqual (2011a). *Criteria for National Vocational Qualifications (NVQs)*. Ofqual/11/4883. Coventry: Office of Qualifications and Examinations Regulation.

Ofqual (2011b). *The 2010 Evaluation of the Qualifications and Credit Framework (QCF) Regulatory Arrangements*. Ofqual/11/4853. Coventry: Office of Qualifications and Examinations Regulation.

Ofqual (2011c). *Criteria for Recognition*. Ofqual/11/4852. Coventry: Office of Qualifications and Examinations Regulation.

Ofqual (2011d). *General Conditions of Recognition*. Ofqual/11/4851. Coventry: Office of Qualifications and Examinations Regulation.

Ofqual (2014a). *Awarding Organisations Operating in the Tier 4 Market: A thematic review*. Ofqual/14/5367. Coventry: Office of Qualifications and Examinations Regulation.

Ofqual (2014b). *A Consultation on Withdrawing the Regulatory Arrangements for the Qualifications and Credit Framework*. Ofqual/14/5490. Coventry: Office of Qualifications and Examinations Regulation.

Ofqual (2014c). *Withdrawing the Regulatory Arrangements for the Qualifications and Credit Framework*. Ofqual/14/5570. Coventry: Office of Qualifications and Examinations Regulation.

Ofqual (2015a). *After the QCF: A new qualifications framework*. Ofqual/15/5652. Coventry: Office of Qualifications and Examinations Regulation.

- Ofqual (2015b). *After the QCF: A new qualifications framework. Decisions on conditions and guidance for the Regulated Qualifications Framework (RQF)*. Ofqual/15/5777. Coventry: Office of Qualifications and Examinations Regulation.
- Ofqual (2016). *Annual Qualifications Market Report England, Wales and Northern Ireland 2014/15 Academic Year*. Ofqual/16/6077. Coventry: Office of Qualifications and Examinations Regulation.
- Ofqual (2018a). *Corporate Plan 2018–21*. Ofqual/18/6360. Coventry: Office of Qualifications and Examinations Regulation.
- Ofqual (2018b). *Strengthening Vocational and Technical Qualifications*. Ofqual/18/6459/1. Coventry: Office of Qualifications and Examinations Regulation.
- Ofqual (2018c). *Implementation of Technical Qualifications. Consultation on rules and guidance for Technical Qualifications*. Ofqual/18/6422/3. Coventry: Office of Qualifications and Examinations Regulation.
- Ofqual (2019). *Essential Digital Skills Conditions and Requirements*. Ofqual/19/6500/3. Coventry: Office of Qualifications and Examinations Regulation.
- Ofqual (2020). *Regulating Performance Table Qualifications: Decisions on Ofqual's approach to regulating the Technical Award qualifications listed in school and college performance tables by the Department for Education*. Ofqual/20/6589/1. Coventry: Office of Qualifications and Examinations Regulation.
- Ofqual (2023). *Level 3 Alternative Academic Qualifications and Technical Occupation Qualifications: Qualification Level Conditions*. Ofqual/22/6967/3. Coventry: Office of Qualifications and Examinations Regulation.
- Ofsted (1993). *GNVQs in Schools: The introduction of General National Vocational Qualifications 1992*. London: HMSO.
- Ofsted (1994). *GNVQs in Schools 1993/94: Quality and standards of General National Vocational Qualifications*. London: HMSO.
- Ofsted (1998). *Standards of Achievement in Advanced GNVQs in Sixth Forms: 1997*. London: The Stationery Office.
- Ofsted (1999). *The Annual Report of Her Majesty's Chief Inspector of Schools: Standards and quality in education 1997/98*. London: The Stationery Office.
- Ofsted (2004). *Vocational A levels: the first two years*. HMI 2146. London: Ofsted.
- Ormeil, C. (1992). Behavioural objectives revisited. *Educational Research*, 34(1), 23–33.

- Ormell, C.P. (1974). Bloom's taxonomy and the objectives of education. *Educational Research*, 17(1), 3–18.
- Parkes, D. (1994). "Competence" and Context: Sketching the British Scene. *European Journal of Vocational Training*, 1, 23–28.
- Pearce, A.G. (1975). TEC and the Tech. *Education & Training (Education + Training)*, 17(3), 54–55.
- Pearce, J. (1972). *School Examinations*. London: Collier-Macmillan Publishers.
- Pearce, J. (1978). The death of essay – English under BEC. *Journal of Further and Higher Education*, 2(1), 3–10.
- Pedley, F.H. (1964). *A Parent's Guide to Examinations: From primary school to university*. London: Pergamon Press Ltd.
- Peters, A.J. (1967). *British Further Education: A critical textbook*. Oxford: Pergamon Press Ltd.
- Pollard, E., Hadjivassiliou, K., Swift, S., & Green, M. (2017). *Credit Transfer in Higher Education: A review of the literature*. London: Department for Education.
- Popham, W.J. (1978). Domain specification strategies. In R.A. Berk (Ed.). *Criterion-Referenced Measurement: The state of the art* (pp.15–31). Baltimore, MD: John Hopkins University Press.
- Popham, W.J. (1994). The instructional consequences of criterion-referenced clarity. *Educational Measurement: Issues and Practice*, 13(4), 15–18, 30.
- Porter, A. C., Smithson, J., Blank, R., & Zeidner, T. (2007). Alignment as a teacher variable. *Applied Measurement in Education*, 20(1), 27–51.
- PricewaterhouseCoopers (2005). *The Market for Qualifications in the UK*. London: PricewaterhouseCoopers LLP.
- Pring, R. (1971). Bloom's Taxonomy: A philosophical critique (2). *Cambridge Journal of Education*, 1(2), 83–91.
- Pring, R.A. (1995). *Closing the Gap: Liberal education and vocational preparation*. London: Hodder and Stoughton Educational.
- Purcell, J. (2001). National Vocational Qualifications and competence-based assessment for technicians – from sound principles to dogma. *Education + Training*, 4(1), 30–39.
- Pye Tait Consulting (2014). *Analysis of the Consultation on Withdrawing the Regulatory Arrangements for the Qualifications and Credit Framework*. Ofqual/14/5571. Coventry: Office of Qualifications and Examinations Regulation.

- Pye, M.E., Pye, K.C., & Wisby, E. (2004). *Apprenticeships in the UK – their Design, Development and Implementation*. North Yorkshire: Pye Tait Limited.
- QCA (Qualifications and Curriculum Authority) (1998). *Internal Verification of NVQs: A guide for internal verifiers*. QCA/98/164. London: Qualifications and Curriculum Authority.
- QCA (Qualifications and Curriculum Authority) (1999a). *Improving NVQs*. London: Qualifications and Curriculum Authority.
- QCA (Qualifications and Curriculum Authority) (1999b). *Developing an Assessment Strategy for NVQs and SVQs*. QCA/99/396. London: Qualifications and Curriculum Authority.
- QCA (Qualifications and Curriculum Authority) (1999c). *Qualifications 16–19: A guide to the changes resulting from the Qualifying for Success consultation*. QCA/99/393. London: Qualifications and Curriculum Authority.
- QCA (Qualifications and Curriculum Authority) (2000). *Arrangements for the Statutory Regulation of External Qualifications in England, Wales and Northern Ireland*. QCA/00/589. London: Qualifications and Curriculum Authority.
- QCA (Qualifications and Curriculum Authority) (2001). *NVQ Code of Practice*. QCA/02/875. London: Qualifications and Curriculum Authority.
- QCA (Qualifications and Curriculum Authority) (2003). *GNVQ Withdrawal. Phase one report: Features of successor provision and implementation issues*. London: Qualifications and Curriculum Authority.
- QCA (Qualifications and Curriculum Authority) (2004a). *Arrangements for the Statutory Regulation of External Qualifications in England, Wales and Northern Ireland*. QCA/04/1293. London: Qualifications and Curriculum Authority.
- QCA (Qualifications and Curriculum Authority) (2004b). *Evaluation of Technical Certificates in Modern Apprenticeships (February 2004)*. London: Qualifications and Curriculum Authority.
- QCA (Qualifications and Curriculum Authority) (2004c). *A Framework for Achievement. Recognising qualifications and skills in the 21st century. Qualifications and Curriculum Authority Stakeholder consultation: December 2004*. QCA/04/1459. London: Qualifications and Curriculum Authority.
- QCA (Qualifications and Curriculum Authority) (2005). *Comparability Study of Assessment Practice: BTEC National qualifications in business, engineering and media*. QCA/05/1681. London: Qualifications and Curriculum Authority.

- QCA (Qualifications and Curriculum Authority) (2006). *NVQ Code of Practice*. QCA/06/2888. London: Qualifications and Curriculum Authority.
- QCA (Qualifications and Curriculum Authority) (2009). *The Quality and Standardisation of Units in the Qualifications and Credit Framework*. QCA/09/4281. London: Qualifications and Curriculum Authority.
- QCA (Qualifications and Curriculum Authority) (undated). *The Future of NVQs. Report of internal working group, circa 2001*. Unpublished.
- QCDA (Qualifications and Curriculum Development Agency) (2010a). *Make All Your Learning Count: A learner's guide to the Qualifications and Credit Framework*. QCDA/10/4623. Coventry: Qualifications and Curriculum Development Agency.
- QCDA (Qualifications and Curriculum Development Agency) (2010b). *The Qualifications and Credit Framework (QCF) and Higher Education: Working together to benefit learners*. QCDA/10/4612. Coventry: Qualifications and Curriculum Development Agency.
- QCDA (Qualifications and Curriculum Development Agency) (2010c). *Key Questions about Assessment in the QCF*. QCDA/10/5202. Coventry: Qualifications and Curriculum Development Agency.
- QCDA (Qualifications and Curriculum Development Agency) (2010d). *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.
- QCDA (Qualifications and Curriculum Development Agency) (2010e). *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.
- Raggatt, P. & Unwin, L. (1991). *Change and Intervention: Vocational education and training*. London: The Falmer Press.
- Raggatt, P. & Williams, S. (1999). *Government, Markets and Vocational Qualifications: An anatomy of policy*. London: Falmer Press.
- 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.
- Raine, N. (1984). BTEC is launched. *Education + Training*, 26(3), 74–74.
- Richard, D. (2012). *Richard Review of Apprenticeships*. Bootle: School for Startups.
- Riches, N.J. (1980). TEC physics. *Physics Education*, 15(6), 364–367.

- Roberts, P. (1988). The control of the further education curriculum. *The Vocational Aspect of Education*, 40(105), 21–26.
- Sainsbury, D. (2016). *Report of the Independent Panel on Technical Education*. London: Department for Business, Innovation & Skills and Department for Education.
- Schools Council (1965). *Change and Response. The first year's work: October 1964 – September 1965*. London: HMSO.
- Schools Council (1968). *Sixth Form Examining Methods*. Working Paper No. 20. London: HMSO.
- Schools Council (1970). *Changes in School Science Teaching*. Schools Council Curriculum Bulletin 3. London: Evans Methuen Educational.
- Schools Council (1972). *16–19: Growth and Response. 1. Curricular bases*. Schools Council Working Paper 45. London: Evans Methuen Educational.
- Schools Council (1973). *Assessment of Attainment in Sixth-form Science*. Schools Council Examinations Bulletin 27. London: Evans Methuen Educational.
- Sellars, J. (1977). Introduction of courses leading to BEC awards. *Journal of Further and Higher Education*, 1(3), 14–16.
- Shackleton, J.R. & Walsh, S. (1995). The UK's National Vocational Qualifications: The story so far. *Journal of European Industrial Training*, 19(11), 14–27.
- Sharp, P. (1998) The beginnings of GNVQs: An analysis of key determining events and factors. *Journal of Education and Work*, 11(3), 293–311.
- Sharp, P. (1999). *The Development of the Vocational Curriculum for 16–19 Year Olds in Colleges and Schools, 1979–1995*. Leeds: School of Education, University of Leeds.
- Sheldrake, J. & Vickerstaff, S. (1987). *The History of Industrial Training in Britain*. Aldershot: Avebury.
- Shepard, L.A. (1997). *Measuring Achievement: What does it mean to test for robust understanding?* Princeton, NJ: Educational Testing Service.
- Shorrocks-Taylor, D. (1999). *National Testing: past, present and future*. Leicester: British Psychological Society Books.
- Skills Commission (2009). *Progression Through Apprenticeships: The final report of the Skills Commission's Inquiry into apprenticeships*. London: The Skills Commission.
- Skinner, B.F. (1958). Teaching Machines. *Science*, 128(3330), 969–977.
- Smith, G. (1994). Ten years of BTEC. *Journal of the Society of Dyers and Colourists*, 110(3), 93–94.

- Smithers, A. (1993). *All Our Futures: Britain's education revolution*. London: Channel Four Television.
- Smithers, A. (1997). A critique of NVQs and GNVQs. In S. Tomlinson (Ed.). *Education 14–19: Critical Perspectives* (pp.55–70). London: Athlone Press.
- Sockett, H. (1971). Bloom's Taxonomy: A philosophical critique (1). *Cambridge Journal of Education*, 1(1), 16–25.
- Stanley, J. (2015). Learning Outcomes: From policy discourse to practice. *European Journal of Education*, 50(4), 404–419.
- Stanton, G. & Bailey, B. (2004). Fit for purpose? Sixty years of VET policy in England. In G. Hayward & S. James (Eds.). *Balancing the Skills Equation: Key issues and challenges for policy and practice*. Bristol: Bristol University Press, Policy Press.
- Stanton, G. (1995). Quality in qualifications and learning programmes. In J. Burke (Ed.). *Outcomes, Learning and the Curriculum: Implications for NVQ's, GNVQ's and other qualifications* (pp.142–154). London: The Falmer Press.
- Stanton, G. (2012). *Course Design and Pedagogy in Adult Vocational Education: Where are we now and where have we come from?* CAVTL 1st Commissioners' Meeting, 27 June 2012, paper 6. Unpublished.
- Stanton, G., Morris, A., & Norrington, J. (2015). What goes on in colleges? Curriculum and qualifications. In A. Hodgson (Ed.). *Coming of Age for FE? Reflections on the past and future role of further education colleges in the UK* (pp.68–88). London: Institute of Education Press.
- Steadman, S. (1995). The assessment of outcomes. In J. Burke (Ed.). *Outcomes, Learning and the Curriculum: Implications for NVQ's, GNVQ's and other qualifications* (pp.201–212). London: The Falmer Press.
- Stenhouse, L. (1975). *An Introduction to Curriculum Research and Development*. London: Heinemann Educational Books Ltd.
- Stevens, P. (1989). Four models of integration in BTEC engineering courses. *Journal of Further and Higher Education*, 13(2), 85–91.
- Stevens, P. (1993). *City and Guilds of London Institute: A short history, 1878–1992*. London: City and Guilds of London Institute.
- Stratton, N. (1986). Recording achievement: The City and Guilds experience. In P. Broadfoot (Ed.). *Profiles and Records of Achievement: A review of issues and practice* (pp.109–126). East Sussex: Holt, Rinehart and Winston Ltd.
- Sutch, T., Zanini, N., & Benton, T. (2015). *A Level Reform: Implications for subject uptake*. Cambridge: Cambridge Assessment.
- Sutherland, M. (1988). *Theory of Education*. Essex: Longman Group UK Limited.

- TA (Training Agency) (1988a). *Development of Assessable Standards for National Certification. Guidance note 1: A code of practice and a development model*. Sheffield: The Training Agency.
- TA (Training Agency) (1988b). *Development of Assessable Standards for National Certification. Guidance note 3: The definition of competences and performance criteria*. Sheffield: The Training Agency.
- TA (Training Agency) (1988c). *Development of Assessable Standards for National Certification. Guidance note 4: The characteristics of units of competence*. Sheffield: Training Agency.
- TA (Training Agency) (1989a). *Development of Assessable Standards for National Certification. Guidance note 2: Developing standards by reference to functions*. Sheffield: Training Agency.
- TA (Training Agency) (1989b). *Development of Assessable Standards for National Certification. Guidance note 5: Assessment of competence*. Sheffield: Training Agency.
- TA (Training Agency) (1989c). *Development of Assessable Standards for National Certification. Guidance note 6: Verification or monitoring of assessment practices*. Sheffield: Training Agency.
- Tarrant, T. (2000). What is wrong with competence? *Journal of Further and Higher Education*, 24(1), 77–83.
- Tattersall, K. (2007). A brief history of policies, practices and issues relating to comparability. In P.E. Newton, J. Baird, H. Goldstein, H. Patrick, & P. Tymms (Eds.). *Techniques for Monitoring the Comparability of Examination Standards* (pp.43–91). London: Qualifications and Curriculum Authority.
- Taylor, H.T. & Beaumont, T. (1967). Backwards from Crowther? Engineering exams since 1961. *Technical Education (Education + Training)*, 9(8), 350–353.
- TGAT (Task Group on Assessment and Testing) (1988). *Report of the Task Group on Assessment and Testing*. London: Department of Education and Science and the Welsh Office.
- Tomlinson (2004). *14–19 Curriculum and Qualifications Reform: Final report of the Working Group on 14–19 Reform*. DfE-0991-2004. London: Department for Education and Skills.
- 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.

- Tuxworth, E. (1989). Competence based education and training: Background and origins. In J. Burke (Ed.). *Competency Based Education and Training* (pp.10–25). Sussex: The Falmer Press.
- Tyler, R.W. (1931). A generalized technique for conducting achievement tests. *Educational Research Bulletin*, 10(8), 199–208.
- Tyler, R.W. (1932a). Measuring the results of college instruction. *Educational Research Bulletin*, 11(10), 253–260.
- Tyler, R.W. (1932b). Improving test materials in the social studies. *Educational Research Bulletin*, 11(14), 373–379.
- Tyler, R.W. (1936). Needed research in the field of tests and examinations. *Educational Research Bulletin*, 15(6), 151–158.
- Tyler, R.W. (1949). *Basic Principles of Curriculum and Instruction*. London: The University of Chicago Press Ltd. London.
- UCLES (1972). *Syllabuses: Science Subjects 1974–5. General Certificate of Education (Ordinary Level)*. Cambridge: University of Cambridge Local Examinations Syndicate.
- UCLES (undated). *Syllabuses: Science Subjects 1957 and 1958. General Certificate of Education (Ordinary Level)*. Cambridge: University of Cambridge Local Examinations Syndicate.
- UKCES (UK Commission for Employment and Skills) (2011a). *NOS Strategy 2010-2020. Revised strategy: June 2011*. London: UK Commission for Employment and Skills.
- UKCES (UK Commission for Employment and Skills) (2011b). *National Occupational Standards Quality Criteria with Explanatory Notes. Revised: June 2011*. London: UK Commission for Employment and Skills.
- UKCES (UK Commission for Employment and Skills) (2014). *[Untitled: Statement of Intent for National Occupational Standards.]* London: UK Commission for Employment and Skills.
- Unwin, L. (1999). *Jungle Trekking: Vocational courses and qualifications for young people*. Skills Task Force Research Paper 5. Sheffield: Centre for Research in Post-Compulsory Education and Training, University of Sheffield.
- Unwin, L., Fuller, A., Turbin, J., & Young, M. (2004). *What Determines The Impact Of Vocational Qualifications? A literature review*. Research report RR522. London: Department for Education and Skills.
- Wallis, B.C. (1927). *The Technique of Examining Children: A quest for capacity*. London: Macmillan and Co. Limited.

- Ward, L. (1996). Moves to boost training take-up. [The Times Education Supplement Magazine](#). 12 January.
- Watson, J.B. (1925). *Behaviorism*. New York: W.W. Norton & Company Inc.
- Webb, N. L. (2007). Issues related to judging the alignment of curriculum standards and assessments. *Applied Measurement in Education*, 20(1), 7–25.
- Wesson, A.J. (1983a). Behaviourally defined objectives: A critique. Part one. *The Vocational Aspect of Education*, 35(91), 51–58.
- Wesson, A.J. (1983b). Behaviourally defined objectives: A critique. Part two. *The Vocational Aspect of Education*, 35(92), 75–79.
- West, J. (2004). *Dreams and Nightmares: The NVQ experience*. CLMS Working Paper No 45. Leicester: Centre for Labour Market Studies, University of Leicester.
- Wheatley, D.E. (1959). 'City and Guilds' examinations. *The Vocational Aspect of Education*, 11(22), 31–59.
- Wheatley, D.E. (1976). *Apprenticeships in the United Kingdom*. Social Policy Series no. 30. Brussels, Luxembourg: Office for Official Publications of the European Communities.
- 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.
- Whitehead, N. (2013). *Review of Adult Vocational Qualifications in England*. London: UK Commission for Employment and Skills.
- Williams, S. & Raggatt, P. (1996). Competence-based vocational training policy: A case study of the decollectivisation of employment relations. *Journal of Vocational Education and Training*, 48(4), 315–332.
- Williams, S. & Raggatt, P. (1998). Contextualising public policy in vocational education and training: The origins of competence-based vocational qualifications policy in the UK. *Journal of Education and Work*, 11(3), 275–292.
- Williams, S. (1999). Policy failure in vocational education and training: the introduction of National Vocational Qualifications (1986–1990). *Education + Training*, 41(5), 216–226.
- Wilson, K.D. (1983). TEC, BEC and the theory of curriculum change. *Journal of Further and Higher Education*, 7(2), 29–36.
- Wilson, P. (2010). *Big Idea, Small Steps: The making of credit-based qualifications*. Leicester: National Institute of Adult Continuing Education.

- Winch, C. (2011). Skill – a concept manufactured in England? In M. Brockmann, L. Clarke, & C. Winch (Eds.). *Knowledge, Skills and Competence in the European Labour Market: What's in a qualification* (pp.85–101)? Oxford: Routledge.
- Winch, C. (2020). Learning outcomes, progression, and qualifications: Considerations for Vocational and Higher Education in the United Kingdom. In C. Arnold, M. Wilson, J. Bridge, M.C. Lennon (Eds.). *Learning Outcomes, Academic Credit, and Student Mobility* (pp.167–197). Kingston, Canada: McGill-Queen's University Press.
- Winch, C. (2023). Learning outcomes: The long goodbye: Vocational qualifications in the 21st century. *European Educational Research Journal*, 22(1), 20–38.
- Wolf, A. (1989). Can competence and knowledge mix? In J. Burke (Ed.). *Competency Based Education and Training* (pp.39–53). Sussex: The Falmer Press.
- Wolf, A. (1990). Unwrapping knowledge and understanding from standards of competence. In H. Black & A. Wolf (Eds.). *Knowledge and Competence: Current issues in training and education* (pp.31–38). London: Careers and Occupational Information Centre.
- Wolf, A. (1993). *Assessment Issues and Problems in a Criterion-Based System*. London: Further Education Unit.
- Wolf, A. (1995). *Competence-Based Assessment*. Maidenhead: Open University Press.
- 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. (2011). *Review of Vocational Education – The Wolf Report*. London: Department for Education.
- Wolf, A., Burgess, R., Stott, H., & Veasey, J. (1994). *GNVQ Assessment Review Project. Final Report*. London: Employment Department.
- Wolf, A., Scharaschkin, A., Meade, M., & Pettitt, A. (1994). *GNVQs 1993–94: A national survey report. An interim report of a joint project. The evolution of GNVQs: enrolment and delivery patterns and their policy implications*. London: Further Education Unit.
- Young, M. & Allais, S. (2009). Conceptualizing the role of qualifications in education reform. In S. Allais, D. Raffe, & M. Young (Eds.). *Researching NQFs: some conceptual issues* (pp.5–22). Geneva: International Labour Office.

- Young, M. & Allais, S. (2011). The shift to outcomes based frameworks: Key problems from a critical perspective. *Magazin erwachsenenbildung.at*, 14, 02-1-03-10.
- Young, M. (2008). *Bringing Knowledge Back In: From social constructivism to social realism in the sociology of education*. Oxford: Routledge.
- Young, M. (2011). National Vocational Qualifications in the United Kingdom: Their origins and legacy. *Journal of Education and Work*, 24(3-4), 259-282.



© Crown Copyright 2024

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated.

[View the Open Government Licence at The National Archives website](#) or write to:
Information Policy Team, The National Archives, Kew, London TW9 4DU

Published by:

ofqual

2nd floor, 1 Friargate
Station Square
Coventry
CV1 2GN

0300 303 3344

public.enquiries@ofqual.gov.uk

www.gov.uk/ofqual