Cambridge Assessment

Research *Matters*

Issue 30 / *Autumn 2020*

Cambridge Assessment

Proud to be part of the University of Cambridge

Established over 150 years ago, Cambridge Assessment operates and manages the University's three exam boards and carries out leading-edge academic and operational research on assessment in education. We are a not-for-profit organisation.

Citation

Articles in this publication should be cited using the following example for article 1: Cooke, G. (2020). A New Cambridge Assessment Archive Collection Exploring Cambridge English Exams in Germany and England in 1938. *Research Matters: A Cambridge Assessment publication*, *30*, 2–6.

Credits

Reviewers: Nicky Rushton, Irenka Suto, Victoria Crisp, Lucy Chambers, Research Division, Cambridge Assessment
Editorial and production management: Anouk Peigne, Research Division, Cambridge Assessment
Additional proofreading: David Beauchamp, Hannah Williams, Research Division, Cambridge Assessment
Cover image: John Foxx Images
Design: George Hammond
Print management: Canon Business Services

All details are correct at the time of going to print in October 2020

If you would like to comment on any of the articles in this issue, please contact Tom Bramley, Director, Research Division. Email: ResearchDivision@cambridgeassessment.org.uk

This and all previous issues of *Research Matters* are available to download, in full and by individual article, from our website: www.cambridgeassessment.org.uk/research-matters

Research *Matters* / 30

A CAMBRIDGE ASSESSMENT PUBLICATION

Foreword

Anyone reading this copy of *Research Matters* in the future might do well to look at the publication date. At the time of compilation, the Covid-19 pandemic remained a disruptive force in education. The offices of Cambridge Assessment were sparsely populated, and with the suspension of public examinations domestically and internationally, the researchers of the organisation were fully deployed on directly supporting the alternative arrangements which were put in place. Urgently assembled models became controversial and problematic in their implementation, and researchers from Cambridge Assessment worked hard to understand and mitigate the limitations of both the models and the awarding strategy which superseded them. This comprised a massive programme of work, full of novel tasks, to an exacting schedule. But amongst this urgent and vital contingent work, we felt it important to continue to do some of the things which we would have done, such as publish this edition. Although changes in life and education have been effected by the pandemic—and some of them may remain permanent—we must not forget the things which we worried about before it hit, including the curriculum and assessment matters covered in the articles here. For sure things will change, but measurement accuracy, high-quality learning and a sense of our past—witness the extraordinary and excellent article on the 1938 Archive holdings—all will determine the quality of our post-Covid education arrangements.

Tim Oates, CBE *Group Director, Assessment Research and Development*

Editorial

It is perhaps helpful to put into perspective the current disruption to high-stakes school examinations in the UK by considering past turmoils. The first article by our group archivist Gillian Cooke is a departure from the usual assessment research we publish in *Research Matters*, but gives a fascinating historical glimpse of exam board activity in 1938 in the months leading up to the start of the Second World War. Although the article was drafted before the current pandemic took hold, her description of the collection of papers as revealing "hopes, fears, ignorance, frustration, compassion, misplaced faith in authorities and a steely defence of examination standards" may give us pause for reflection!

In the second article, Jo Ireland and Melissa Mouthaan describe some of the metaphors used to understand curriculum design (spirals, networks, webs), and some of the arguments about which are most useful in different fields of knowledge.

Staying with the theme of curricula, the third article by Sinéad Fitzsimons and colleagues describes the high-level principles that should be considered when developing curricula for learners in emergency situations where normal educational provision is disrupted by (for example) war or natural disasters.

While much research has considered the differences between on-screen and paper assessment from the point of view of the test taker, less has considered the question writers. The fourth article by Vicki Crisp and Stuart Shaw gives an account of a detailed investigation of the experiences of question writers in writing and reviewing questions in an on-screen environment.

The final article by Irenka Suto and colleagues shows how different taxonomies of skills and knowledge developed for general academic contexts can be evaluated and deployed in more applied contexts. They argue that taxonomies are an underused tool that could help improve validity in curriculum and assessment design.

Tom Bramley *Director, Research Division*

A New Cambridge Assessment Archive Collection Exploring Cambridge English Exams in Germany and England in 1938

Gillian Cooke Cambridge Assessment Archives and Heritage

While the Archives of Cambridge Assessment sit within the Research Division and can be drawn on for traditional assessment research studies, this work represents a different type of research into the historical operations of the organisation. It focuses on a collection of correspondence and papers from 1938, recently acquired by Cambridge Assessment.

In 1938, Europe was on the cusp of the Second World War. The National Socialist dictatorship in Germany was aggressively pursuing policies to create an Aryan German Empire while many Cambridge English candidates at that time in Germany were Jewish. With 75th anniversary commemorations since the liberation of war time death camps in Europe (in 1945) still fresh, this is a timely look back at the role Cambridge Assessment, then the University of Cambridge Local Examinations Syndicate (generally referred to as UCLES or "the Syndicate"), played at the beginning of this period to keep the examinations flowing and to meet the demands of prospective candidates.

Archival documents are the raw data of the historian. When the historian pores over the primary source as an unexplored gem, it is not just the contents of the document under scrutiny, but the context, provenance and diplomatic qualities of the physical document that shape the interpretation. The interest is often in the minutiae, as archives offer the researcher information about topics that may never have been considered.[1] By cataloguing, the archivist aims to present original documents in a clear, objective way, but here I have prepared a form of extended catalogue as an essay which draws on each document in the collection. I have also linked some of the documents within the small collection, and with others in related archive collections, to create a narrative. In this, I have gone beyond the role of archivist, but the principle intention, to highlight the documents for further historical research, remains the same.

On 8 March 1938, G. H. Gretton, an English teacher in Hamburg, wrote to Jack Roach at the University of Cambridge Local Examinations Syndicate. Following up on his query about an English *Précis* examination, Gretton thanked Roach for clarification and apologised for his delayed reply. He regretfully declined a lunch invitation at Peterhouse in Cambridge, then told Roach of his decision to leave Germany—he needed a job for the following September—and, he revealed, "*I don't like the temper of recent exchanges.*"[2] With its references to Cambridge, to specific aspects of English examinations and to life in Germany under the regime of the Nazi dictatorship, this one letter captures the essence of a remarkable collection of papers which were presented to Cambridge Assessment in 2018.

The collection has survived through serendipity; rescued from a skip by Bill Shephard, a successor to Jack Roach pictured in Figure 1, it fell into the hands of his family after his death and has now been added to the catalogued collection of Jack Roach papers.[3] All the documents were written in 1938. Here I will consider what these papers tell us about Roach and the Cambridge Exams during this turbulent period, how Roach responded to different groups, and how his attitudes changed during the year. The collection broadly falls into two parts; one part consists of letters exchanged between Roach at the Local Exams Syndicate and correspondents in Germany, most of whom are Jewish teachers; the other part relates to syllabus development and the promotion of Cambridge English exams, and includes correspondence with Roach's UK contacts. Most of the correspondents are unidentified beyond the information contained in the collection.[4]

Figure 1: Jack Roach in the 1940s. Cambridge Assessment Archives Ref: PP/JOR 10/2.

Jack Roach was a Modern Linguist and Europhile who became Assistant Secretary (deputy head) of UCLES in 1925. He inherited a tiny new exam called the Certificate of Proficiency in English (CPE), which he believed passionately should expand into an untapped market for English Language learning in Europe. The Certificate of Proficiency did not start out as a purely English qualification. It was launched in 1913 as part of Modern Languages and Religious Knowledge proficiency qualifications for over 18-year-olds but, while the French, German and Religious Studies candidates enjoyed initial success, the English papers proved to be fiendishly difficult and carried a very low pass rate. Eager to increase the accessibility and reach of the English qualification, Roach

persuaded the Syndicate to lower the exam fee and remove the particularly troublesome phonetics paper from the syllabus in 1932.[5] He then convinced it to allow him grace leave to travel around Europe in the summer of 1937, to *"carpetbag for my examinations in Europe."*[6] After returning home he began to revise the CPE, develop the Lower Certificate of Proficiency in English as a sister qualification, and write a visionary *'Memorandum'* to encourage the co-ordination of facilities for foreign language students in England, thereby promoting and increasing the candidature of Cambridge English qualifications throughout Europe.

The earlier Roach collection includes a log of his tour of Europe in summer 1937 and visits to two Jewish schools in Berlin: Dr Goldschmidt's European Jewish school, which prepared candidates for School Certificate and English Proficiency (and had grown considerably since becoming a centre); and the smaller private *Waldschule Kaliski*, a Zionist day school. Roach enjoyed *"tea and long conversation"* at Dr Goldschmidt's school, but notes a *"slight uneasiness [at Walsdschule Kaliski] over the fact that Dr Goldschmidt's school is the only centre."* Centre status, he observed, represented a particular *"privilege"* to German teachers and parents *"because in German recognised schools the staff themselves conduct the examinations, mark the papers and hold the oral tests."*[7] Schools recognised as centres by the Syndicate were printed on an annual centre list which included Dr Leonore Goldschmidt's school in 1938.[8]

It should not, therefore, have surprised Roach to receive a request, the following January, for centre status from *Waldschule Kaliski*, although it is remarkable to see the next chapter of this story revealed in this new collection of documents. In his reply, Roach counsels against exaggerating the establishment of an examination centre *"into a privilege which in fact it is not,"* adding: *"we cannot give you a statement to the effect that you have been authorised by the Syndicate to prepare candidates for the examinations, since no such authority is required."* But Dr Selver, the Director, is persistent; having successfully prepared four candidates the previous month, he writes of a further 12 candidates who can enter without seeking permission from German government bodies. Roach's response shows some irritation, but eventually he agrees to the school becoming a closed centre for one year *"if, and only if, we receive the approval of the German authorities."* Roach then writes to the British Consulate in Berlin, advising them of a prospective new centre; *"We shall be glad to know whether, in the event of the centre being forced, you could nominate a suitable supervisor"*, he asks, revealing ambiguous deference in his use of the word *"forced"*.[9] Later correspondence indicates that the school duly became the second centre to provide regular courses of preparation for the CPE in Berlin.[10]

Accessibility to books is an issue Roach raised in 1937 on his European tour; *"the whole question of text-books is very difficult for these people"* he observed after visiting the two Jewish schools in Berlin. By 1938, these difficulties had escalated to examination fees and currency transfer. In February, the Rev L. G. Forrest, Chaplain to the Church of England in Hamburg, asked to transfer exam fees through British Embassy in Berlin, but Roach regretfully declined; *"I am therefore afraid that the nuisance must continue"*.[11] But the *"nuisance"* grew; in May, the Director of *Waldschule Kaliski* hinted at a new and complex governmental process, and in October the new director, Dr Jacob, accompanied his late entry with the plea: *"I sincerely ask you in this case not to insist on the extra fee of 5 shillings, as we would have to approach again the 'Devisenstelle' (the State Foreign Exchange Board) which we would very much like to avoid."* Roach does waive the extra fee *"in the circumstances"* but adds *"on this present occasion"* to the same sentence, just to emphasise his discomfort at this irregularity.[12]

Adherence to the standards and regulations is a common theme that runs through the correspondence with frequent references to deadlines, regulations and fees. We cannot know exactly what any of the correspondents knew of political circumstances at any given time during 1938, and although Roach's stance becomes less dogmatic as the year advances, it is unclear if this did much to facilitate access to the qualifications. In February, Roach referred Dr Goldschmidt to the revised regulations, paragraph 15, to refuse her appeal for a candidate to take only a partial retake, but he did waive the late entry fee.[13] In the same month, he provided Hamburg-based Gretton with exacting instructions about the *Précis* exam.[14] In September, Roach offered to advance the introduction of the new Lower Certificate exam to meet *Waldschule Kaliski's* needs, but the gesture elicits a flat response from Dr Jacob; *"even the month of June… seems in view of the present situation still too late."*[15] Indeed, while the Lower Certificate was introduced in June 1939, there are no candidates from Germany included in these first pass lists.[16]

Four of the German correspondents mention the urgent need of English qualifications for migration, citing high prospective candidate figures; Dr Goldschmidt promises 70 candidates, while Dr Jacob claims that 100 are being prepared for examination within the year.[17] However, while the subsequent pass lists for the CPE include many Berlin candidates, and candidates from these schools, they fall short of these figures. This suggests that the numbers could have been exaggerated, events prevented the candidates from taking the exams, or, that many candidates failed to pass.[18]

There is a clear understanding that English qualifications were needed for German Jews to emigrate for work or study in English-speaking countries, and, as the year progresses, this pressure intensifies. In January, Dr Selver at *Waldschule Kaliski* tells Roach that currency issues prevent study in England, but that *"most of our pupils intend going [to America] sooner or later"* to which Roach confidently offers to *"make the situation clear… if you have any specific difficulty with the American authorities."*[19]

In September 1938, having succeeded Dr Selver (who himself has emigrated to the US), Dr Jacob tells Roach of new requirements for *"Jewish physicians…to give up their positions on 1st October."* *"The circumstances,"* he stresses, *"are at present extraordinary and more powerful than our real educational considerations."* This is Dr Jacob's first letter as the new head of *Waldschule Kaliski*, and he makes the most of the opportunity to bring Roach on side. After updating Roach on the fortunes of the successful candidates who have reached America or Jerusalem, he goes on to *"heartily thank the Cambridge examinations syndicate, and especially you, dear Mr Roach… who gave us the possibility for this examination to be held here at the School."*[20]

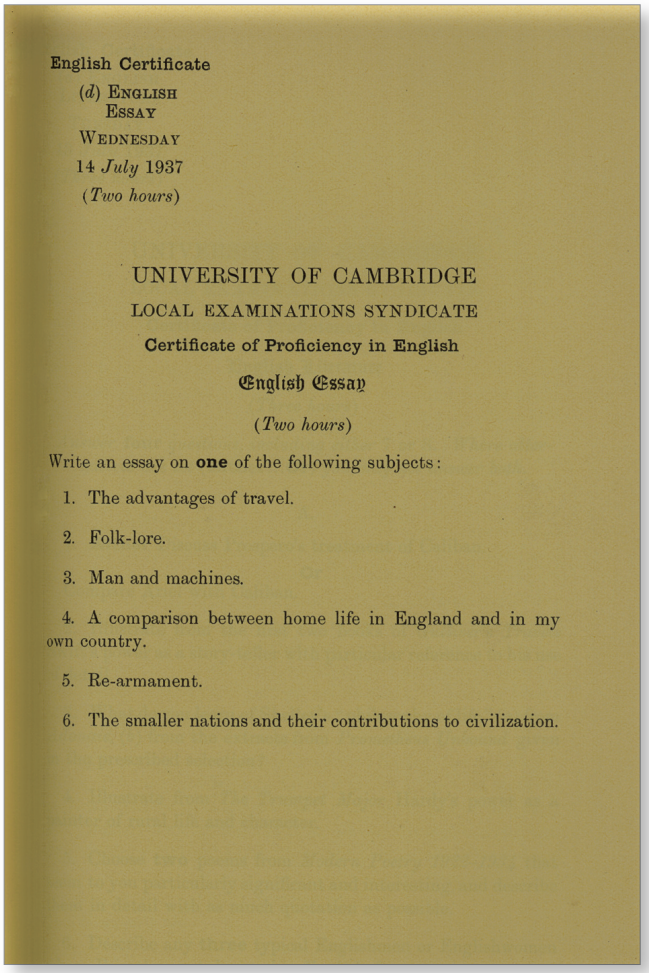
By this time, Dr Goldschmidt has asked Roach to support her appeal to the University State of New York for recognition of the CPE. *"It is very important for immigrants into USA, specially doctors of medicine, if they can pass their linguistic examination as a recognised one by the American authorities before their depart from Berlin"*, she explains. The response she received from Horace Field at the Bureau of Qualifying Certificates, University State of New York, dismisses the *"so-called Proficiency in English examination offered by your institution"* in a tone that must have piqued Roach, and he takes up the cause. He asks for the CPE to be given equivalent status to the state English admission exam, citing recognition

of the CPE by the Universities of Oxford, Cambridge and Scotland, but his pitch is misplaced, and Field remains unimpressed.[21]

The ensuing correspondence shows that the American authorities required a monitored comparison of the standard between the two exams; they ask for *"at least a dozen"* successful CPE candidates to come forward to take the US admissions exam. Roach's realisation that this is extremely unlikely, his sensitive explanation to Dr Jacob of the difficulties confronting them,[22] and his letter to the Director of the University of the State of New York on 19 October, reveal a sharp awareness of difficult circumstances. He explains that the candidates in question are *"Jewish doctors and others in Germany [who] cannot guarantee their movements for long in advance, and they even find it difficult to obtain examination fees."* He offers the assistance of representatives of British Jews and of the Cambridge English Oral Examiner in Berlin, that they may *"guide us all towards a solution to this problem."* He references individual risk and the need to save time and money, and he forwards copies to all the Syndicate's Local Secretaries and the US Ambassador in Germany with a warning *"lest undue hopes should be raised."*

Ultimately the US authorities stand firm, their conditions are not met, and Roach concedes defeat. The day following *Kristallnacht*, when Jewish schools throughout Germany were ransacked, Roach writes to his German correspondents: *"there is no immediate prospect of obtaining recognition for the CPE as exempting from the test in English for foreigners who are candidates for admission to professional licensing examinations in the state of New York……Jewish doctors must judge for themselves whether they should act as pioneers in this matter in order to help others, but the Syndicate cannot give them any advice on this matter or hold out any particular hope of recognition."*[23] While Roach fails to gain recognition of the CPE for US professional licensing examinations, it does show his greater awareness of the political situation over that of his US correspondents, a sensitivity to the plight of Jews in these professions, and a degree of tenacity.

Moving to the second part of the collection, Roach's correspondence with his contacts at home deals with the development of the Certificate of Proficiency and the new Lower Certificate. His letters to H. L. Ellis at the City of London College bristle with some of the ideas he included, five months later, in his thirteen-page *Memorandum*. They discuss textbooks for a controversial new paper for social life, customs and government. *"Some books, however acceptable in themselves, might be unpopular in, say, Germany and Italy"*[24] writes Ellis. The suitability of examination content is also raised by A. L. Jeavons, Principal of the Marlborough London County Council Evening Institute in March 1938. His observations on the Summer 1937 papers are accurate and he draws uncomfortable analogies from the *"prophetic utterance"* of the prose passage to the annexation of Austria by Germany, earlier that month. He reserves his harshest criticism for the examiner who *"made the oral test an opportunity to air his views on Germany's coup d'état in Austria, anti-semitism, and Dr. Niemoller"* (an outspoken Christian critic of the Nazi Regime). *"Personally, I agree with much that he said"*, writes Jeavons, frankly *"but it was surely bad taste to give an impression of antagonism, particularly in circumstances which do not call for the expression of any political opinion whatsoever."* Roach's response is tellingly unapologetic, he explains that the question paper was set long before recent events in Austria and defends the examiner's *"sane British point of view"*, but he also invites suggestions for future content of the CPE, and the exam questions in 1939 are distinctly more benign.[25]



English Certificate

(*d*) English Essay

Wednesday

14 *July* 1937

(*Two hours*)

UNIVERSITY OF CAMBRIDGE

LOCAL EXAMINATIONS SYNDICATE

Certificate of Proficiency in English

English Essay

(*Two hours*)

Write an essay on **one** of the following subjects:

1. The advantages of travel.
2. Folk-lore.
3. Man and machines.
4. A comparison between home life in England and in my own country.
5. Re-armament.
6. The smaller nations and their contributions to civilization.

Figure 2: Certificate of Proficiency in English, English Essay Question Paper, July 1937. Cambridge Assessment Archives Ref: Bound Volume, 1937.

Although Roach is the principal advocate of the Cambridge English exams, it is clear from the collection that he is also heavily involved in the overseas candidature of the School Certificate and Higher School Certificate exams as well, and the collection includes references to standards between these qualifications. Jeavons made a strong case for disparity between the literature standards of the School Certificate and the CPE, and Roach admits: *"we are perhaps handicapped by the fact that we have to provide for candidates in different countries and with very varying facilities for study."*[26] Roach was mindful of overseas centres which prepared candidates for both School Certificate and Proficiency, noting in his 1937 log that Dr Goldschmidt employed eight teachers for the School Certificate.[27] He also referenced the standard of School Certificate English in correspondence with the US authorities at New York State University.[28] Development of the new Lower Certificate syllabus complicated these discussions; while the introduction of a new lower level English exam was welcomed by Sir Stephen Gaselee at the Foreign Office, and by representatives in Baghdad, Bucharest and Tripoli, the Syndic S. W. Grose was more cautious. He warned of the limitations of a vocabulary list set at around 1500 *"essential"* words from the Oxford English Course,[29] and, in September, Roach was forced to address Jacob's misinterpretation of the standard of the new qualification at *Waldschule Kaliski*.[30]

References to costs for European students are not confined to Roach's German correspondents and are peppered throughout the collection. The undated draft regulations for the Lower Certificate reference the affordability of the new exam, at ten shillings,[31] and Ellis suggested

obtaining text books more cheaply from abroad for preparation of the new British Life and Institutions Paper for the Certificate of Proficiency.[32] In his August *Memorandum*, Roach declares *"our great difficulty for foreign students is the price of English books"*. He also advocates the use of public libraries and inter-library loan services,[33] and raises the possibility of collaboration with the Oxford University Press.[34]

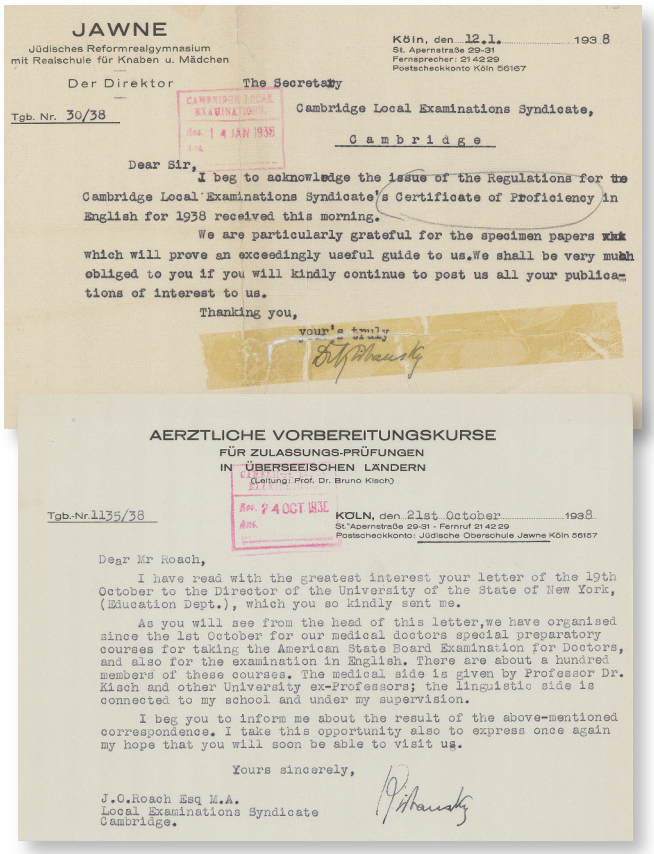
In August 1938, Roach produced two papers on *the co-ordination of facilities for foreign students in England* in which he is critical of the English attitude to foreign language students. In contrast to the formal civility of his correspondence with German centres, here Roach is less restrained. The papers are accompanied by a long letter to Sir Eugene Ramsden, MP and member of the British Council: *"There is much room for improvement, even in the treatment of foreign newspaper correspondents by persons in more or less high places"* he argues, and, with startling foresight: *"The war is on, democracy against totalitarianism and it must be so!"*[35]

Addressed to the Department of Intelligence at the Board of Education, the *Memorandum* and its *Addendum* are less emotive but nevertheless promote radical ideas. They advocate state sponsorship of English language candidates, recommending incentives for cultural integration, including the dissemination of literature about Britain, subsidised concert tickets and provision of transport. Roach proposes lecture tours dealing with democratic government and the British way of life, the registration of all foreign students entering the UK and the exchange of labour for learning (which brings him into conflict with the Ministry of Labour on the definition of an *"au pair"*[36]). He recommends support committees for cases such as *"the girl paying three guineas a week in a very lower middle class family with bad cooking and no social life"* but is wary of developing a *"reputation of the Home Office and its police surveillance."* His criticism is not limited to the Home Office; to achieve his aims he invites collaboration with the Foreign Office and the British Council while taking a swipe at both, by drawing uncomplimentary parallels between the British Council and bureaucratic government departments. His particular irritation is a government grant of £25,000 a year which he believes the British Council should distribute more evenly, and his view is supported by R. Howland, his Syndicate ally, who admits: *"The British Council seems the obvious body to do something useful."* The *Memorandum Addendum* is written in response to a *Times* article by Professor Trevelyan where Roach predicts financial benefits to the economy in welcoming foreign students to England. Here, he challenges prospective English examination candidates to drop *"an examination phobia about their set books"* and embrace the new Life and Institutions paper to *"make a serious study of our parliamentary democracy, our justice, our local government, our social service."* Roach references costs and affordability, and unashamedly advocates a *"discreet twist given to their reading"* without reference to the suitability of texts which preoccupied him and Ellis in March. Throughout this long, cultural call-to-arms, there is just one reservation: *"I am assuming that we shall surmount the present crisis and that there will be no general war"* he adds, in parenthesis, to Sir Eugene.[37]

Roach's proposals may be considered foolish or admirable but his intentions, towards the English reception of foreign students, and promotion of Cambridge English examinations, are well placed and clear. The aim of the paper is ultimately to gain Board of Education support for the Cambridge English exams, along the same lines as support given to School Certificate candidates at approved schools: *"Is it too much to suggest that the taxpayer might properly pay £1 a head towards the examination fees of candidates taking either of the two approved proficiency examinations in the United Kingdom?"* he muses.[38]

It is easy to dismiss Roach's papers as fantasy, particularly as they were so eclipsed by the outbreak of war the following year, but it is likely that the political momentum itself fuelled the environment for his ideas. The *Memorandum* clearly represents a shift in Roach's attitudes from earlier in the year, but however well-intentioned his stance towards foreign students, he seemed to have been unprepared for the speed of events.

The documents, and lack of them, towards the end of the year give glimpses of difficulties, actions and courage from correspondents which must have informed Roach and the Cambridge Local Exams Syndicate of an escalating political crisis, and undoubtedly contributed to Roach's shifting attitude. In March, following Austria's annexation, Roach queried the return of the unopened examination papers from Vienna.[39] The absence of a reply would have fed his growing unease. In October, Roach learned that *JAWNE*, the *Jűdisches Reformrealgymnasium* in Cologne, had re-invented itself, as shown in Figure 3, when the same director, of the same centre, wrote to him from the *Aerztliche Vorbereitungskurse* on newly headed paper where the word *"Jűdisches"* has been carefully concealed. (Now aptly named for *"special preparatory courses for taking the American State Board Examination for Doctors, and also for the examination in English"*).[40] In the autumn, Roach reassured Triebig that candidates *"are quite at liberty to prepare privately for the examinations"*[41] and Dobson, also in Berlin, that the *"rumours of a new lower certificate"* are indeed true.[42] There is no further correspondence from Dr Jacob or Dr Goldschmidt after October but the final entry for the year is a typescript note, written by Roach on 8th December,



JAWNE
Jüdisches Reformrealgymnasium
mit Realschule für Knaben u. Mädchen

Köln, den 12.1. 1938
St. Apernstraße 29-31
Fernsprecher: 214229
Postscheckkonto Köln 56167

Der Direktor

Tgb. Nr. 30/38

The Secretary
Cambridge Local Examinations Syndicate,
Cambridge

CAMBRIDGE LOCAL EXAMINATIONS Rec'd 14 JAN 1938

Dear Sir,

I beg to acknowledge the issue of the Regulations for the Cambridge Local Examinations Syndicate's Certificate of Proficiency in English for 1938 received this morning.

We are particularly grateful for the specimen papers which will prove an exceedingly useful guide to us. We shall be very much obliged to you if you will kindly continue to post us all your publications of interest to us.

Thanking you,

your's truly

AERZTLICHE VORBEREITUNGSKURSE
FÜR ZULASSUNGS-PRÜFUNGEN
IN ÜBERSEEISCHEN LANDERN
(Leitung: Prof. Dr. Bruno Kisch)

Tgb-Nr.1135/38

Rec'd 24 OCT 1938

KÖLN, den 21st October 1938
St. Apernstraße 29-31 · Fernruf 214229
Postscheckkonto: Jüdische Oberschule Jawne Köln 56167

Dear Mr Roach,

I have read with the greatest interest your letter of the 19th October to the Director of the University of the State of New York, (Education Dept.), which you so kindly sent me.

As you will see from the head of this letter, we have organised since the 1st October for our medical doctors special preparatory courses for taking the American State Board Examination for Doctors, and also for the examination in English. There are about a hundred members of these courses. The medical side is given by Professor Dr. Kisch and other University ex-Professors; the linguistic side is connected to my school and under my supervision.

I beg you to inform me about the result of the above-mentioned correspondence. I take this opportunity also to express once again my hope that you will soon be able to visit us.

Yours sincerely,

J.O.Roach Esq M.A.
Local Examinations Syndicate
Cambridge.

Figure 3: Correspondence from the Jewish English Centre in Cologne, which changed its name to Aerztliche Vorbereitungskurse in 1938, thereby hiding its Jewish affiliations, January and October 1938. Cambridge Assessment Archives Ref: PP/JOR 11/10 and PP/JOR 11/13.

referencing an encounter with Leonore Goldschmidt in Cambridge, *"here to make efforts to transfer batches of her pupils to England."* He notes that he has enlisted her help to contact two Cambridge local examination officials in Berlin to explain that if relations are broken with them *"this would in no sense be due to a feeling that they had done less than their duty to us while also fulfilling their loyal duty to the service which employs them."*[43]

This quotation could well apply to Roach, whose sense of duty to prospective candidates, and the Local Examinations Syndicate, is evident throughout this collection. And, despite an overwhelming sense of foreboding, a glance at the pass lists shows just how much was achieved. The steady increase in candidate numbers peaks at the end of 1938 and includes candidates from Dr Goldschmidt's school and *Waldschule Kaliski*.[44] Indeed, the correspondence from these centres is testimony that some successful candidates did make it out of the country; furthermore, Dr Goldschmidt made her own extraordinary achievement in relocating her Jewish school to Folkstone in 1939.[45]

This collection is just a snapshot of Roach's work for UCLES in one exceptional year, and this study highlights just some aspects of the extraordinary communications it contains. Despite the formality of the business correspondence and papers, it is clear that attitudes, including Roach's own, shifted as the year advanced. But this collection has far more to give; set against a backdrop of Third Reich rule in Germany it reveals hopes, fears, ignorance, frustration, compassion, misplaced faith in authorities and a steely defence of examination standards.

Nearly a century later, there are overtones of these experiences in some of Cambridge Assessment's work today. Second World War experiences may be seared into our organisational history, but conflict and community displacement continue, with growing demand to provide monitored standards of educational assessment to displaced learners. As our organisation has grown, and understanding of humanitarian crises matures, we are increasingly able to develop imaginative and appropriate responses. The most recent and far reaching of these, the UNICEF Learning Passport Project of 2019,[46] represents a university-wide collaboration, with other sponsors, to develop Literacy, Mathematics and Science curricula to displaced learners over a range of ages and backgrounds. Its scale and complexity is incomparable to our response to the circumstances of 1938 but, in the context of identifying displaced learners and responding to need, it remains helpful for Research colleagues in Cambridge Assessment to be aware of a real and fallible historical precedent.

The following references relate mainly to documents held in Cambridge Assessment Archives. The reference PP/JOR 11 is the prefix for documents in the new Roach collection in 1938.
For details of how to access this archive collection, please contact archives@cambridgeassessment.org.uk

1. Powell, Nia. Obituary of A. D. Carr, Archivist and Historian; Archives and Records, 2019, Vol 40 No 3.
2. PP/JOR 11/8; Gretton to Roach 8th March.
3. Cambridge Assessment Archives: PP/JOR.
4. For background information about UCLES examinations at this time, see *Examining the World*, edited by Sandra Raban, CUP, 2008.
5. PP/JOR 1/1ai; Bound Volumes, 1931 & 1932.
6. PP/JOR 2/1; Report on European Visit by Roach 1937.
7. PP/JOR 2/1; Ibid.
8. Bound Volume, 1938.
9. PP/JOR 11/7; Waldschule Kaliski and Roach 26th Jan, 3rd May & 16th May.
10. PP/JOR 11/11; Roach to Triebeg, 23rd Sept.
11. PP/JOR 11/6; Forrest and Roach 10th March.
12. PP/JOR 11/7; Waldschule Kaliski and Roach 11th May & 3rd Oct.
13. PP/JOR 11/9; Goldschmidt and Roach 17th Feb.
14. PP/JOR 11/8; Roach to Gretton 24th Feb.
15. PP/JOR 11/7; Waldschule Kaliski and Roach 3rd Oct.
16. P/EFL 13/2; LCE June 1939 Pass List.
17. PP/JOR 11/10; Recognition of exams in US 3rd Sept & 11/7; 7th Sept.
18. P/EFL 13/1; Dec '38 – 5 & 20 candidates; Mar '39 – 12 & 13; Jun '39 – 3 & 0 from Dr Goldschmidt's School & Waldschule Kaliski respectively.
19. PP/JOR 11/7; Waldschule Kaliski and Roach 22nd & 26th Jan.
20. PP/JOR 11/7; Waldschule Kaliski to Roach 7th Sept.
21. PP/JOR 11/10; Recognition of exams in US 16th Aug, 3rd Sept, 20th Sept, 3rd Oct & 6th Oct.
22. PP/JOR 11/7; Waldschule Kaliski and Roach 20th Sept.
23. PP/JOR 11/10; Recognition of exams in US 19th Oct & 10th Nov.
24. PP/JOR 11/3; Ellis and Roach 10th March.
25. PP/JOR 11/4; Jeavons and Roach 30th March & 1st April.
26. PP/JOR/11/4; Ibid 1st April.
27. PP/JOR 2/1; Report on European Visit by Roach 1937.
28. PP/JOR 11/10; Recognition of exams in US 20th Sept.
29. PP/JOR 11/2; Proposal for Lower Certificate 26th April & 6th June.
30. PP/JOR 11/7; Roach to Waldschule Kaliski 20th Sept.
31. PP JOR 11/2; Proposal for Lower Certificate undated.
32. PP/JOR 11/3; Ellis and Roach 10th March to 8th April.
33. PP/JOR 11/1; Paper on Co-ordination of Facilities for Foreign Students in England Roach and correspondents August.
34. PP/JOR 11/2; Proposal to Lower Certificate undated.
35. PP/JOR 11/1; Ibid 30th August.
36. PP/JOR 11/5; Ministry of Labour and Roach 29th Aug & 5th Sept.
37. PP/JOR 11/1; Paper on Co-ordination of Facilities for Foreign Students in England; Roach and correspondants August.
38. PP/JOR 11/1; Ibid.
39. PP/JOR 11/14; Roach to Captain Taylor, British Consulate, 29th March.
40. PP/JOR 11/13 & 10; 12th Jan and 21st Oct.
41. PP/JOR 11/11; Triebig and Roach 22nd Sept.
42. PP/JOR 11/12; Dobson and Roach 5th Oct.
43. PP/JOR 11/10; Recognition of exams in US 8th Dec.
44. P/EFL 13/1; Pass Lists March 1938–June 1939 include 82 candidates from Germany of which 30 Dr Goldschmidt & 33 Waldschule Kaliski.
45. Dr Gertrud H Thompson: leonoregoldschmidt.com
46. Cambridge University Press & Cambridge Assessment. (2019). *The Learning Passport Research and Recommendations Report: Summary of Findings*. Cambridge, UK: Cambridge University Press & Cambridge Assessment. https://www.cambridgeassessment.org.uk/Images/577273-the-learning-passport-research-and-recommendations-report..pdf

Perspectives on curriculum design: comparing the spiral and the network models

Jo Ireland and **Melissa Mouthaan** Research Division

Introduction

Does one approach fit all when it comes to curriculum design? In debates on curriculum design, educators have argued that a curriculum model should take into account the differing knowledge structures of different subjects. Subjects such as Mathematics and Science are generally defined as well-structured knowledge domains, characterised by a linearity in learning objectives, and well-defined and predictable learning outcomes. Less structured subjects such as the arts and humanities could, however, benefit from models that encompass a different approach to learning. Two competing perspectives on curriculum design have emerged: the spiral model developed by Bruner (1960) and non-linear models based on processes of learning in different knowledge domains (Efland, 1995, 2000; Yang, 2000). Research on curriculum design has tended to focus on the needs of Science, Technology, Engineering and Mathematics (STEM) subjects. Many alternative models to the spiral have come from arts-based disciplines, in particular visual arts.[1]

This article contributes to the ongoing debate about curriculum design in different subjects. It details the key characteristics of Bruner's spiral model, and presents the main arguments made in favour of adopting flexible and non-linear curriculum models in specific subjects. We discuss a number of alternatives to the spiral model and analyse the relative strengths and weaknesses of these different approaches. The conclusion offers a discussion of implications of our findings for further research in curriculum design.

Background: the spiral curriculum

Bruner (1960) developed the spiral curriculum model by drawing on the way that concepts and knowledge are structured in the physical sciences. This was considered to be equally applicable to the arts and humanities. Learning is visualised as a spiral upwards from basic to advanced concepts, with topics being revisited at increasing levels of complexity as the spiral loops round. The process of reinforcement in learning is a key feature of the spiral curriculum. Each return visit has additional objectives and presents fresh learning opportunities. In a spiral curriculum, attention is paid to both the scope and sequence of topics. Bruner felt it was important that learners obtain the most "fundamental understanding" of a subject by having a solid grasp of the underlying principles of that subject (Bruner, 1960, p.31). In particular, he speculated that if learners were introduced to specific topics or skills without a connection to core principles in the broader field of knowledge, they would:

a) be unable to generalise from what has been learned and apply this in other scenarios;
b) find little 'reward' in terms of intellectual excitement; and
c) be more likely to forget what they have learned, if this knowledge is not structurally organised in terms of principles and ideas (Bruner, 1960, pp.31–32).

Knowledge structures

At the heart of the spiral curriculum theory is Bruner's assertion that "any subject can be taught effectively in some intellectually honest form to any child at any stage of development" (Bruner, 1960, p.33). This signals his firm belief that the spiral curriculum could apply to all subjects. Yet, this assumption has also formed the key contention that advocates of alternative models have brought against Bruner's model.

The assumption that it is possible and/or desirable to extrapolate from Science to other subjects has led to criticism of the spiral model, particularly from those concerned with the arts (Efland, 1995). Drawing on findings in cognitive research, these critiques have highlighted that specific knowledge domains are structured differently from the STEM topics on which Bruner based his spiral model. They have argued that the relative ill-structuredness of some domains is a poor fit with processes of learning captured by the spiral. In this sense, the spiral curriculum is found to have disadvantages in its application to less-structured knowledge domains.

Feltovich et al. (1993) used the term 'ill-structured' to describe domains which require a learner to synthesise many different concepts, and patterns of concepts, on a case-by-case basis. This type of knowledge is found in many fields including law, literary criticism, history and philosophy: any subject where there is an "absence of rules or generalizations that apply to numerous cases" (Efland, 2002, p.84). Finding "a key idea around which to organize instruction" is also less evident in subjects that rely less on the study of over-arching principles, and more on the in-depth study of specific cases (Efland, 2000, p.278; 2002). Learning through understanding laws, axioms or theorems, where problems have a single correct solution, is more common in well-structured subjects (Short, 1995; 1998). This consistency is less common in some of the social sciences, humanities and arts (Alexander et al., 1991; Short, 1998). A subject's underlying structure, it is argued, has key implications for learning within that subject.

Likewise, the theory of vertical and horizontal discourse (Bernstein, 1999) described different forms of knowledge as hierarchical (e.g., science) or horizontal (e.g., humanities). Hierarchical knowledge structures appear to be "motivated towards greater and greater integrating propositions, operating at more and more abstract levels" (Bernstein, 1999, p.162). Horizontal knowledge structures "consist of a series of specialised languages with specialised modes of interrogation and criteria for the construction and circulation of texts" (p.161).

1. In comparison, curriculum models for subjects such as Literature, languages and the humanities have not received the same level of scrutiny.

Trying to apply a single curriculum model to these two types of knowledge structures presents obvious problems. The spiral model seems to fit more naturally with the hierarchical knowledge structure of the sciences and the move towards abstract ideas.

Cognitive researchers have argued that the process of tailoring the complexity of ideas to early stages of learning constitute 'simplification strategies' (Efland, 2000) or a 'reductive bias' in the spiral model (Feltovich et al., 1993; Spiro et al., 1988). They argued that the reduction of complexity that occurs in the spiral curriculum has implications, the most notable being a single representation of ideas at the expense of multiple representations. Instead, they argued that learners should be encouraged to study ideas and concepts in all of their complexity.

Reduced complexity favours single representations (e.g., a single schema, organisational logic, line of argument, or analogy). The use of simplification in instruction is a helpful tool, particularly in early stage learning, which enables a learner to interpret a new concept using existing knowledge. However, as Spiro et al. (1988) argued, singular representations carry a risk of missing the many aspects of a complex concept, while learners may also fail to develop diversified ways of thinking. Using the example of learning in medicine, studies have argued that singular representations can form simplification strategies in learning that are obstacles for developing in-depth, advanced learning strategies at later stages (Feltovich et al., 1993; Spiro et al., 1988). Feltovich et al. (1993) described biomedicine as an ill-structured domain where "the linkage between surface features of cases and applicable concepts is irregular and rich, relational indexing and categorisations are not only particularly important but also particularly difficult for the learner to construct" (p.202). Therefore, one problem with the notion that the spiral curriculum begins with simple concepts and progresses to mastery is that it fails to recognise that for ill-structured domains the spiral model can lead to misconceptions in early learning, which persist into advanced study. However, while reductiveness is intended to make knowledge acquisition easier, Efland (2000) argued that it may lead to students not understanding what is being taught, and struggling to relate the knowledge to their own lives.

How has the spiral model been applied?

Science

Comparing science education in China to the United States, Su et al. (1995) found that Chinese students who were taught science via a spiral curriculum developed good theoretical knowledge and basic skills, while the United States took a 'layered' interdisciplinary curriculum approach where students developed good factual recall. A comparison of the science curricula of each country (Herr, 2007) showed that China taught Biology, Chemistry and Physics at each grade level between Grade 7 and 12—a 'vertical and spiral' model. The United States covered a broader range of subjects, including Environmental Science and Zoology, and taught Biology only at Grade 10, Chemistry only at Grade 11 and Physics only at Grade 12. Laboratory work in the United States was interdisciplinary. However, these comparisons do not tell us much about the relative merit of each approach, given cultural and societal differences between the countries' education systems.

Medicine

Harden and Stamper (1999) related how the spiral model informed the University of Dundee medical course. Year 1 focuses on structure, function and behaviour. In Years 2 and 3, students revisit these concepts when studying abnormal structure, function and behaviour. This knowledge is then related to clinical practice in Years 4 and 5, and finally the theory is put into practice in students' pre-registration year. In the case of medicine, it seems logical that a solid theoretical base must be the starting point before students commence practice, and the outline given by Harden and Stamper fits Bruner's theory. However, as we highlighted earlier, medicine is also characterised by its ill-structuredness.

Mathematics

In the United States, the spiral curriculum is the mostly widely used structure for school mathematics (Seely, 2009; Snider, 2004). The effectiveness of the spiral curriculum in the United States has been questioned, mainly due to perceived poor performance compared to other countries in Trends in International Mathematics and Science Study (TIMSS; Snider, 2004). Schmidt et al. (2005) created charts of common content standards for Mathematics by examining the curriculum structures of six top-performing jurisdictions (as measured by TIMSS). They found that increasing mathematical complexity was introduced as students progressed through school years with some topics forming 'buttresses' across multiple school years.[2] This was characterised as a 'staggered spiral'. They then compared this data with United States content standards and found that the United States featured longer duration of topics, with the majority of topics being covered across all the US Grades 1–8. They referred to this situation as the 'mile-wide inch-deep curriculum'. Despite the differing approaches, all countries were described as using the spiral curriculum.

Music

Swanwick (1979) proposed a set of hierarchical music learning objectives, which was later expanded to produce a model of musical development (Swanwick & Tillman, 1986). Taking Piaget's child development theories as a starting point, Swanwick and Tillman (1986) applied the concepts of *mastery* (control of sound materials), *imitation* (expressive character/accommodation) and *imaginative play* (structural relationships/assimilation) to a music learning context. They observed children aged between 3 and 9 years old in Music lessons and found that the compositions of children followed this sequence. Furthermore, with reference to other studies, they felt able to tie the stages of development to particular age groups.

Table 1: Stages of musical development

Piagetian concept	*Description of musical enactment*	*Stage*
Mastery	From sensory exploration to manipulative skills.	Ages 0–4
Imitation	Personal and idiosyncratic expression to socially shared vernacular conventions/reproducibility.	Ages 4–9
Imaginative	Speculative composition and attention to formal musical devices.	Ages 10–15
Meta-cognition	Self-awareness of thought processes and feelings in response to music.	Age 15+

2. One limitation of the analysis was that the aggregated data was not representative of any single one of the countries' complete curriculum.

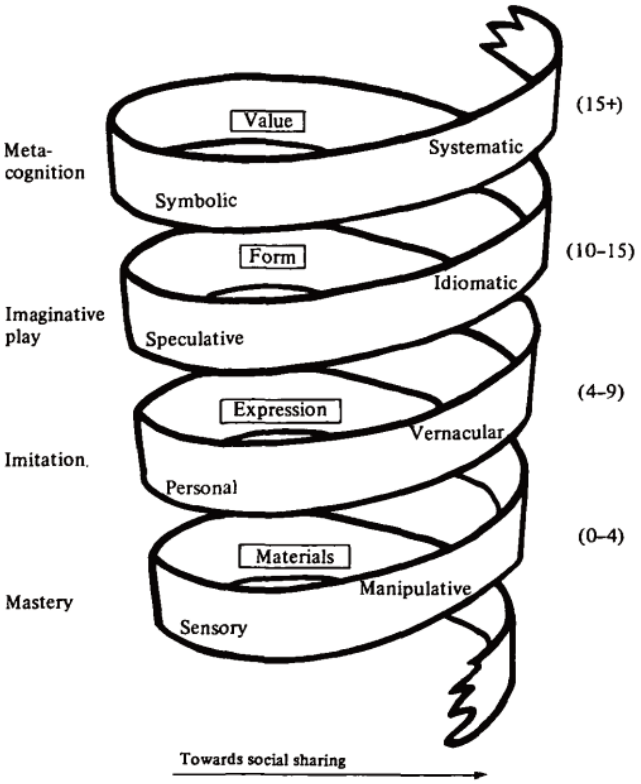


Figure 1: Spiral model of musical development (Swanwick & Tillman, 1986, p.331). © Cambridge University Press 1986, reproduced with permission.

This development model was visualised as a spiral (Figure 1).

The spiral curriculum's central concept of revisiting stages or topics is incorporated into Swanwick and Tillman's model. For example, they state that "the shift from sensory exploration towards manipulative skills […] is an on-going concern at any stage of development" (1986, p.320). Furthermore, they proposed that development stages or topics were revisited even by experienced musicians, giving the example of the exploratory behaviour entailed in familiarisation with a new instrument (1986, pp.336–337).

Swanwick later reflected on the spiral model and conceded that the Swanwick and Tillman spiral (Figure 1) was just one of a number of ways of conceptualising musical development. He argued that as well as linear and quantitative, progression in Music could also be seen as layered and qualitative, with issues such as critical judgement and audience response forming part of the model (Swanwick, 2016).

The Manhattanville Music Curriculum Project (MMCP) was an education programme that ran in the 1960s and 1970s in the United States (Moon & Humphreys, 2010). The programme aimed to reform music education and was based on the spiral curriculum model. Elements of music (dynamics, timbre, form, rhythm and pitch) were organised into cycles which repeated with increasing complexity. In the programme, students were presented with problems and they then composed and performed their responses, followed by an evaluation stage. This is an interesting contrast to the Swanwick and Tillman spiral, because the MMCP model refers directly to the subject content rather than the development stages.

Art

As evidence for his hierarchical theory of musical development, Swanwick cited the work of Hargreaves and Galton (1992). This is a hierarchical model of artistic development which describes progress in cognitive aesthetic development, drawing, writing, singing, musical representation, melodic perception and musical composition. Five phases of development are proposed: Presymbolic (ages 0–2), Figural (2–5), Schematic (5–8), Rule Systems (8–15) and Metacognitive (15+). Stages of development then populate the model; for example, the Presymbolic phase of drawing would be demonstrated by scribbling and the Rule System phase of melodic perception by analytic recognition of intervals. While not a spiral model in itself, the sequential progression of development is in keeping with the spiral philosophy.

Discussing the ARTS PROPEL[3] approach, Gardner (1989) believed that, if suitably structured, a spiral model could be valuable in arts curricula in schools. While he discounted the idea of atomistic curricular goals for the arts, he argued that sequential, holistic goals could be included in an arts curriculum and that core concepts such as style, composition and genre were revisited at increasing levels of sophistication. Gardner described these aims as a spiral model.

Advantages and disadvantages of the spiral model

Harden and Stamper (1999) noted a number of advantages to the spiral model:

- Reinforcement—it encourages retention of knowledge;
- Simple to complex—topics are introduced in a controlled way, to enable better understanding;
- Integration with other parts of the curriculum—subject silos are broken down;
- Logical sequence—attention is paid to the sequence of topics at the curriculum design stage;
- Higher level objectives—increasing complexity encourages students to move beyond recall to application of knowledge.

However, educators from different fields have suggested a number of shortcomings in the application of the spiral model. Comparing different approaches to Mathematics curricula, Snider argued that although the intent of the spiral model is to treat each concept with increasing depth at successive grade levels, the "functional result is that students acquire a superficial understanding of math concepts" (2004, p.31). She identified several aspects of the spiral design that may contribute to this:

- Superficial treatment of topics—students may fail to master important concepts as the spiral model promises further opportunities for mastery with subsequent visits;
- Those who do master the concepts are subjected to unnecessary repetition of the content which can be demotivating (Jensen, 1990);
- Topics introduced at an inappropriate rate—concepts are allotted the same amount of time whether easy or difficult to master;
- Minimisation of academic learning time[4]—the rate at which new content is introduced can mean students unsuccessfully grapple with difficult concepts, or lose interest due to a lack of challenge;

3. ARTS PROPEL was a collaborative project implemented in the USA in the 1980s. It sought to describe the competences arts students should display: production, perception and reflection, with learning named as a core concern.
4. Academic learning time is defined as the amount of time students are 'actively, successfully, and productively engaged in learning' (Brodhagen & Gettinger, 2012).

- Insufficient cumulative review—learning reinforcement is hindered if there is too much time between visits to topics.

One notable feature of Bruner's book, *The Process of Education* (1960), is the lack of evidence offered for the efficacy of the spiral approach. The scarcity of literature assessing the impact of spiral curricula suggests that in general the spiral has been implicitly accepted as a key model for curriculum design, with limited critical reflection on its suitability across subjects. The reason for this might be, as Johnston (2012) noted, "because the spiral curriculum is often interwoven with other inquiry-based and constructivist learning approaches, it is often quite difficult to assess the effects of the *curriculum*, rather than the *delivery* of that curriculum" (pp.1–2). Alternatively, the perception of science subjects as well-structured and hierarchical may have led to an assumption that the spiral curriculum is appropriate. Educators in arts-based disciplines have argued that positivism as an influence in the development of, for example, social inquiry within the social sciences, has favoured structured approaches, such as that of the spiral model (Efland, 2002; Sullivan, 1989). The influence of positivism in research may therefore also be a reason for the lack of scrutiny in applying the spiral approach to art disciplines.

Other studies consider the structured sequencing of learning objectives within the spiral model to be a disadvantage, rather than a strength. Venable (1998) used the example of teaching art criticism to argue that inflexible sequencing may prevent learners from engaging with a topic in depth, as it creates a situation where certain outcomes are intended to be pre-cast, discouraging connections to other areas. Short (1995) similarly argued that in-depth thinking requires not only conceptual and factual knowledge, but also "cognitive flexibility [in order] to see numerous relationships between the two" (p.167). Both studies perceived these types of horizontal connections to other topics to be limited by the structured sequencing within the spiral.

Non-linear curriculum models: network and web models

Most non-linear models that have been proposed as an alternative to the spiral take the form of a 'network' or a 'web'. Webs and networks put less emphasis on linear progression in a knowledge domain and the development of discrete skills, and more emphasis on 'meaning-centred' approaches (Slattery, 2006, p.116; Efland, 2002). In these models, the learners' grasp of the interconnectedness of ideas and the importance of transfer of learning between contexts is emphasised. For instance, Perkins (1989) considered that "understanding something entails appreciating how it is 'placed' in a web of relationships that give it meaning" (p.114). In these models, the sequencing of learning objectives can be flexible, and learners can participate in the choice of their learning pathway. While web or network models have tended to emerge from art domains, they are by no means the only fields to utilise such models. For example, Cambridge Mathematics (2018) has developed an evidence-based, non-linear framework of mathematics knowledge.

Several studies have highlighted the challenges of curriculum design in ill-structured domains, and have subsequently proposed alternatives. The landscape model was proposed by Spiro et al. (1988) in response to the authors' concerns with the single knowledge representation and the reduction of complexity in the spiral model. They argued that a curriculum landscape must be criss-crossed in many directions to master its complexity, reflecting the emphasis on multiple interpretations. The 'lattice' structure was in turn proposed by Efland (1995, 2000) who drew inspiration from the landscape model. The lattice is an alternative to the spiral model, yet both are described by Efland as geometric forms constituting a representation of three factors:

a) The way knowledge is organised in an individual's knowledge base;
b) The way domains of knowledge are organised; and
c) The way content is arranged for purposes of instruction.

The lattice model specifically allows for the overlapping and interconnecting of ideas. This maintains the inherent complexity of a knowledge domain, and addresses the need for the multiple representations that Efland and his contemporaries have advocated. The complex organisation of the model enables "multiple routes of intellectual travel" between and among overlapping domains of knowledge (Efland 2000). The role of transfer—when the learner grasps common elements between two different ideas or concepts—is therefore facilitated in this model.

Efland acknowledged particular flaws in the lattice model. Specifically, he perceived a risk that the lattice could introduce too much complexity in the early stages of learning, and that it is not constrained by natural boundaries, with the potential to spread outwards. An important distinction between the lattice and the landscape model is that Spiro et al. envisaged a model of a domain awaiting discovery by exploration; in comparison, the lattice was intended to function as a "structure actively undergoing construction as learning progressed" (Efland, 2002, p.100). Addressing concerns about the risk of over complexity in the lattice model, Yang (2000) and Efland (2002) went on to propose models with more clearly defined boundaries.

The city model was proposed in order to reflect a better balance between capturing interconnectedness between ideas, and the need to avoid overwhelming the learner with detail at early stages (Efland, 2002). The model was built on the idea of learning as travel within different parts of a city, involving movement from one domain to another, and is similar to the lattice model in this regard. Efland likened curriculum plans to city plans, and stressed the feature of overlapping sets, where the same facts appear in separate domains of knowledge. Overlapping sets act as 'points of transfer', where learners familiar with knowledge in one domain have a possible entry point to begin their exploration of another domain. In this way, learners might study a painting, and learn about the historical context of the painting or the historical event that the painting depicts. The city model also gives learners agency in the learning process as they can choose the destination of travel. In travelling to unfamiliar domains, the learner can benefit from guidance provided by teachers or other knowledgeable peers who act as mediators (Efland, 2000). Efland noted that learners also have a choice in their 'method' of travel where this choice reflects their cognitive strategies; for instance, taking the underground covers a greater range of territory and is a faster mode of transport, but travelling on foot allows exploration of a topic in greater detail.

A related model is a delineated travel network (Yang, 2000), where this travel network has natural boundaries that the lattice does not. For instance, airlines do not organise flights between all cities as this would be too complex to maintain, but rather certain cities are instituted as hubs or transfer points. In curriculum terms, a hub might consist of a broad theme through which one might reach a variety of related destinations (Efland, 2002, p.103). As with Efland's city model, there are

'connecting points' that learners may revisit throughout their learning, leading to the gradual comprehension of a complex domain, or concept, over time.

Conclusion

This article has outlined the approach captured by Bruner's spiral curriculum model, its main features, and its application in varied subjects. While the spiral model has been widely applied since the 1960s in different contexts, educators from some fields have argued that the spiral is better suited to well-structured subjects. We have outlined alternative curricular models and summarised the arguments in favour of them, noting where these models go beyond the spiral's limitations, but may also have their own limitations.

The literature we have reviewed indicates two distinct views on curriculum design. The spiral's structured approach to the scope and sequencing of learning objectives ensures knowledge outcomes are pre-planned, while also enabling vertical integration within the curriculum as topics are revisited. Repeat visits of topics at increasing levels of complexity, a key feature of the spiral, also places importance on the learners' grasp of core concepts, whereby ideas are built on to achieve mastery. On the other hand, proponents of web or network models argue that learning is not always linear, that simplification strategies in learning are unhelpful, and that connections between concepts are vital for integrated learning. Beyond the vertical integration of topics within a spiral curriculum, advocates of non-linear models argue that there is a need to forge horizontal connections between ideas and knowledge domains. These type of 'network' models advocate a more flexible approach to the sequencing and scope of learning objectives, where learners also have decision-making power in their learning journey.

These ideas are presented in the literature as opposing views, with Mathematics and Science requiring a strict linear and hierarchical approach and arts subjects demanding a non-linear alternative, but in fact we have presented evidence that elements of both views can apply whether a subject is well- or ill-structured. While most of the research into alternatives to the spiral model has originated in arts contexts, our findings suggest that these considerations and conclusions can be applied to well-structured subjects as well as ill-structured, as seen with Cambridge Mathematics (2018). Likewise, some commentators have seen merit in a spiral approach for arts subjects (Swanwick & Tillman, 1986; Gardner, 1989).

While this article has highlighted how non-linear models favour greater horizontal exploration of ideas within a curriculum, the risk of overload onto students' learning in a model defined by a lack of natural boundaries is also a valid concern. Efland acknowledged that the lattice model and the risk of introducing too much complexity at the early stages of learning is an example of a tendency towards over-complexity. At the same time, it appears that little has been done to explore the opportunities within the spiral model for facilitating these types of connections. We have noted that the spiral model has a greater tendency to predetermined knowledge outcomes than the approach adopted by non-linear models. Yet the need for conceptual and factual learning that is found in the spiral model, and the need to understand key ideas, is not altogether absent in alternative models. We consider that bridging these two different perspectives provides an avenue for future work on curriculum design.

References

Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research, 61*(3), 315–343.

Bernstein, B. (1999). Vertical and Horizontal Discourse: An Essay. *British Journal of Sociology of Education, 20*(2), 157–173.

Brodhagen, E. M., & Gettinger, M. (2012). Academic Learning Time. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp.33–36). Springer US.

Bruner, J. S. (1960). *The process of education*. Harvard University Press.

Cambridge Mathematics (2018). *In More Detail: Supplement to March 2018 Framework Update*. https://www.cambridgemaths.org/images/cambridge-mathematics-symposium-2018-framework-supplement.pdf

Efland, A. (1995). The Spiral and the Lattice: Changes in Cognitive Learning Theory with Implications for Art Education. *Studies in Art Education, 36*(3), 134–153.

Efland, A. (2000). The City as Metaphor for Integrated Learning in the Arts. *Studies in Art Education, 41*(3), 276–295.

Efland, A. (2002). *Art and cognition: integrating the visual arts in the curriculum*. Teachers College Press.

Feltovich, P. J., Spiro, R. J., & Coulson, R. L. (1993). Learning, teaching, and testing for complex conceptual understanding. In *Test theory for a new generation of tests*. (pp.181–217). Lawrence Erlbaum Associates, Inc.

Gardner, H. (1989). Zero-Based Arts Education: An Introduction to Arts Propel. *Studies in Art Education, 30*(2), 71–83.

Harden, R. M., & Stamper, N. (1999). What Is a Spiral Curriculum? *Medical Teacher, 21*(2), 141–143.

Hargreaves, D. J., & Galton, M. (1992). Aesthetic Learning: Psychological Theory and Educational Practice. In B. Reimer & R. A. Smith (Eds.), *Yearbook on the Arts in Education*. NSSE.

Herr, N. (2007). *The Sourcebook for Teaching Science*. http://www.csun.edu/science/index.html

Jensen, R. J. (1990). One Point of View: The Spring Is Wound Too Tight on Our Spiral Curriculum. *The Arithmetic Teacher, 38*(1), 4–5.

Johnston, H. (2012). *The Spiral Curriculum*. https://eric.ed.gov/?id=ED538282

Moon, K.-S., & Humphreys, J. T. (2010). The Manhattanville Music Curriculum Program: 1966–1970. *Journal of Historical Research in Music Education, 31*(2), 75–98.

Perkins, D. N. (1989). Art as understanding. In H. Gardner & D. N. Perkins (Eds.) *Art, mind, and education: Research from Project Zero* (pp.111–132). University of Illinois Press.

Schmidt, W. H., Wang, H. C., & McKnight, C. C. (2005). Curriculum Coherence: An Examination of Us Mathematics and Science Content Standards from an International Perspective. *Journal of Curriculum Studies, 37*(5), 525–559.

Seely, C. L. (2009). *Faster Isn't Smarter: Messages About Math, Teaching, and Learning in the 21st Century*. Math Solutions.

Short, G. (1995). Understanding Domain Knowledge for Teaching: Higher-Order Thinking in Pre-Service Art Teacher Specialists. *Studies in Art Education, 36*(3), 154–169.

Short, G. (1998). The High School Studio Curriculum and Art Understanding: An Examination. *Studies in Art Education, 40*(1), 46–65. JSTOR.

Slattery, P. (2006). *Curriculum development in the postmodern era* (2nd ed.). Routledge.

Snider, V. E. (2004). A Comparison of Spiral Versus Strand Curriculum. *Journal of Direct Instruction, 4*(1), 29–39.

Spiro, R., Coulson, R., Feltovich, P., & Andersen, D. (1988). *Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains* (Tech. Rep. No. 441). Center for the Study of Reading, University of Illinois.

Su, X., Goldstein, S., & Su, J. (1995). Science Education Goals and Curriculum Designs in American and Chinese High Schools. *International Review of Education, 41*(5), 371–388.

Sullivan, G. (1989). Curriculum in Art Education: The Uncertainty Principle. *Studies in Art Education, 30*(4), 225–236.

Swanwick, K. (1979). *A Basis for Music Education*. Routledge.

Swanwick, K., & Tillman, J. (1986). The Sequence of Musical Development: A Study of Children's Composition. *British Journal of Music Education, 3*(3), 305–339.

Swanwick, K. (2016). Music Development: Revisiting a Generic Theory *A Developing Discourse in Music Education*. Routledge.

Venable, B. (1998). Questioning the Assumptions behind Art Criticism. *Art Education, 51*(5), 6–9.

Yang, G.-M. (2000). *Exploration of Chinese art using a multimedia CD-Rom: design, mediated experience, and knowledge construction* [Ohio State University, Graduate School]. http://rave.ohiolink.edu/etdc/view?acc_num=osu1488203158826307

Context matters—Adaptation guidance for developing a local curriculum from an international curriculum framework

Sinéad Fitzsimons, Victoria Coleman, Jackie Greatorex Research Division, **Hiba Salem** Faculty of Education, University of Cambridge and **Martin Johnson** Research Division

Colleagues across the University of Cambridge worked alongside UNICEF and Microsoft to develop the Learning Passport (LP).[1] The aim of the LP is to contribute to achieving the UNICEF goal of providing a quality education provision to the over 30 million children and youth worldwide who are unable to access a quality education provision due to disruptions caused by crisis and displacement. This area of education is often referred to as Education in Emergencies (EiE). Education in Emergencies refers to education which takes place in an emergency situation, such as a crisis or disaster which disrupts consistent education provision. The EiE landscape is diverse, with a range of learners, learning environments and facilitators. Developing a universal curriculum or learning programme to be used unilaterally across all EiE contexts would not be a logical or ethical method for providing support (Cambridge Assessment, 2020). Instead, it was decided that a blueprint curriculum framework would be created which would provide a set of minimum concepts and principles, integrated into parsimonious learning sequences. These learning sequences would then serve as knowledge-based blueprints for localised curriculum development across a variety of contexts.

The LP project resulted in a curriculum framework for Mathematics, Science and Literacy (Cambridge Assessment, 2020). Alongside this framework, Adaptation Guidance was also created. The Adaptation Guidance was directed towards curriculum experts that would be responsible for developing a localised curriculum based on the LP framework. Although intended to be used in the EiE context, this curriculum development guidance is relevant to curriculum experts across all educational contexts. With global movements of people consistently increasing in recent decades, the demographic of classrooms is changing in most urban areas and in many rural schools

1. More details are available at https://www.cambridge.org/files/8615/8465/3596/The_Research_and_Recommendations_Report.pdf

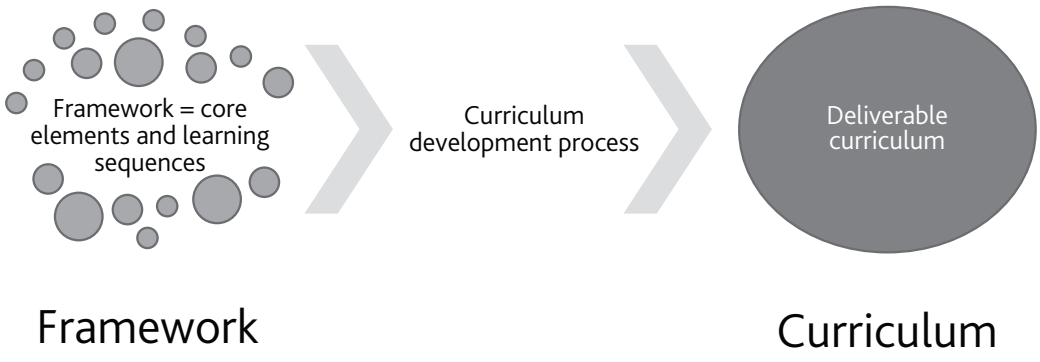


Figure 1: Developing a curriculum from a parsimonious learning framework.

as well (Sugarman, 2015). Currently, in many European classrooms the responsibility for actively including, accommodating and supporting migrant children in schools falls primarily to the teacher (European Commission/EACEA/Eurydice, 2019). However, this is not enough. Nor should the responsibility only be in the hands of teachers. A recent UNESCO report (2018) argues that laws and policies are failing migrant and refugee children and ignoring their needs, especially in terms of education provision. Furthermore, the report attests that provision alone is not enough. The learning environment must adapt and support the specific needs of those on the move. A well-designed curriculum is part of this supportive environment. UNESCO argues that curricula must be inclusive and relevant for learners, including migrant learners. A curriculum that focuses on learner diversity can have a positive ripple effect both within and beyond the classroom walls (UNESCO, 2018).

This article will consider how key guidance areas within the Learning Passport Adaptation Guidance can be applied to wider educational contexts. This will be prefaced by a brief overview of the Learning Passport Framework and the Learning Passport Adaptation Guidance in order to provide further context to the curriculum guidance that emerged.

The Framework

When conceptualising a framework, it is helpful to use the analogy of laying the foundations of a building. In this way, the framework is a guiding structure, or blueprint, for the construction of a curriculum which, importantly, would require crucial localised adaptation to make it relevant to learners in any particular context.

This interplay of centralised control and localised flexibility is an affordance of the concept of a framework. Using the analogy above, centralised control allows a generalised standard (the structural stability of the building) to work with locally relevant features that make the building fit with the surrounding cultural specificity. In terms of a learning programme, the framework provides the underpinning progression structure of important concepts that has generalisable qualities with the ability to transfer across learning contexts.

The framework approach of the LP project allowed the development team to avoid developing a strict universal curriculum programme that would potentially impose a set of knowledge, skills and understandings on all learners without incorporating the localised day-to-day experiences, culture, prior understandings and their desires for the future (Cambridge University Press & Cambridge Assessment, 2020). In this way, as few references as possible were made to specific contexts (e.g., environmental or cultural references) or materials (e.g., devices or tools required to reach understanding). Instead, the framework serves as a blueprint of the essential elements of knowledge-focused content that should be incorporated into a curriculum in order to support quality learning in that subject area. In short, the aim was to create a framework that could serve as a broad outline that covers the minimum requirements of key learning concepts, understandings and principles.

The Adaptation Guidance

The aim of the Adaptation Guidance was to provide a list of considerations that local curriculum developers should reflect upon before they begin their curriculum development process. These guidance areas were developed based on the findings of the Research and Recommendations Report (Cambridge Assessment & Cambridge University Press, 2020) and through consultation with internal and external experts. The Adaptation Guidance was also reviewed by an external review group (ERG) that consisted of over 30 specialists working across the EiE field, including curriculum developers, practitioners, policy developers, and leaders of charities and Non-Governmental Organisations.

The five guidance areas are:

1. Curriculum developers must take into account relevant curriculum and education policies as well as previous learning experiences.
2. Locally adapted curricula should be developed and delivered in the most appropriate language(s) of instruction, after thorough consideration of a variety of factors.
3. Content in locally adapted curricula should be framed so that it is culturally sensitive.
4. Indigenous knowledge should be included in the locally adapted curricula.
5. Locally adapted curricula should support learner well-being, inclusion and success.

These elements focus specifically on curriculum development and are not intended to provide guidance relating to pedagogy, resources and assessment, although these guidance areas have some applicability to these aspects. As each guidance area is reviewed, connections with broader educational contexts beyond that of EiE will be reflected upon. Through this discussion, this article aims to highlight that the challenges that exist within some of the most deprived educational contexts have applicability in supporting quality and equality in education in all educational contexts, including the most affluent.

1. Curriculum developers must take into account relevant curriculum and education policies as well as previous learning experiences

In order to support continued learner development, the contextualisation process must consider the previous educational experiences of learners and educators as well as potential educational pathways that learners may encounter in the future. When developing a curriculum based on the LP framework, the development team should seek to consider, compare and potentially integrate elements of relevant local curriculum. This will allow learners to build on previous understandings and to support them when accessing future education pathways. To aid this process, curriculum mapping and consultation with local curriculum specialists is recommended as it can help curriculum developers identify areas of overlap and potential gaps in previous learning (Elliott, 2011, 2014; Greatorex et al., 2019). However, we recognise that due to the ad-hoc nature of some EiE education provisions, accessing information related to a learner's previous or future educational environment may not be possible. In addition to curriculum documentation, it is also vital that curriculum developers consider education policies and contextual circumstances that may have impacted the learning experiences of different groups of learners, including, but not limited to, restrictions associated with gender, ethnicity or cultural group.

The importance of considering relevant educational policies and previous education experiences that have impacted learners go beyond

the EiE. At present, many migrant children are expected to quickly assimilate into new classrooms and follow a prescribed curriculum in order to succeed (Clark, 2017). Beyond filling in forms explaining their previous education, little is done to investigate the educational experiences these learners have gone through. The process of considering previous learning as well as future education ambitions often falls to the responsibility of the classroom teacher who, after getting to know their learners and reading their learning profiles, accommodates and differentiates the content to support their needs (Clark, 2017). Some curricula allow for flexibility to incorporate relevant content and skills to support migrant learners; however, little is often done to make the curriculum itself more relevant (UNESCO, 2018).

Curriculum developers and regional or school-based curriculum facilitators should take into account student demographics in order to identify relevant curriculum and education policies that may have impacted on previous learning experiences of students. Considering previous educational experiences should not be seen as a beneficial addition, but as a necessary measure to ensure students are appropriately supported and challenged. This is not just impacting a small group of learners. In London alone, it is predicted that there are 254,000 foreign-born children and approximately 107,000 undocumented children who have either arrived in the UK illegally or who were born to undocumented parents (Jolly et al., 2020).

Although considering the previous learning experiences of all students would be difficult from a curriculum development level, considering the make-up of the student demographic in regional areas would allow curriculum developers and facilitators to focus on the larger groups in that area, in order to ensure the curriculum structure, content and expectations cohere with, are flexible to, and build on already established learning. Knowing where learners are likely to move to in their next stage of education allows curriculum designers to ensure that the curriculum progression structure will prepare learners towards reaching this point. In addition, consulting international curricula and curriculum policies can help to highlight discriminatory approaches and practices that are either explicitly or implicitly incorporated into the curriculum content and materials that migrant students interact with (Taylor & Sidhu, 2012).

2. Locally adapted curricula should be developed and delivered in the most appropriate language(s) of instruction, after thorough consideration of a variety of factors

Although the LP framework is presented in English, the curriculum derived from it is meant to be developed and delivered in a language(s) of instruction (LOI) that is most appropriate for its context of use. The decision of which language or languages to choose is worthy of careful consideration.

The use of a learner's mother tongue or native language is important for a number of reasons. Research has shown that learners thrive most when they are taught in a language they understand, as well as a language that will help them to succeed at the time of learning and in the future (Cambridge Assessment & Cambridge University Press, 2020). It is also the case that oracy and literacy development in the mother tongue or native language supports learners in acquiring the communication and understanding skills required to facilitate learning of additional languages with greater ease (Pinnock & Vijayakumar, 2009).

Language choice also links with a rights-based approach to learning (Sandkull, 2005; Capstick & Delaney, 2016). The UN Declaration on the Rights of Minorities (UN General Assembly, 1992) advocates that adequate opportunities should be provided for linguistic minorities to have educational instruction in their mother tongue. Language allows the voice of displaced learners to be expressed. This is even more important where other rights are withheld (Kosonen, 2005) and the lack of access to a learner's native language acts as a form of linguistic discrimination (Romaine, 2013).

If a learner's native tongue or a language that they are competent in cannot be used for instruction, then several other areas of consideration should be reflected upon. For example, contexts such as the Kakuma camp in Kenya, where around 87 languages are spoken (Forsen & Guvatt, 2015), pose significant challenges in selecting a LOI. Such cases are exacerbated where the languages in use are 'distant', meaning that the languages differ greatly in terms of their phonetics, syntax and semantics (Nerbonne & Hinrichs, 2006). If it is not possible to offer the LOI in a language understood by all students, then the strategy of code-switching may help. Code-switching allows learners to move between languages through drawing on the common features of several language systems (Setati & Adler, 2000). To support this strategy, curriculum developers and teachers may use informal language when introducing a new process or concept rather than simply using a nominalised term.

Curriculum development teams should also be wary of issues of language prestige and status. Dearden (2014) reports that there is a general trend towards expansion of English as the LOI because it is believed to provide learners with more future opportunities. However, choosing a 'prestige language' as the LOI when there is a lack of quality teaching, support and resources for that language can result in ineffective pedagogy, inaccurate content delivery and lower-quality materials, which "perpetuat[es] the cycle of educational impoverishment" (Marinotti, 2016, p.5).

LOI choices in education can also be linked to a legacy of colonialism. For example, the linguistic divisions in Cameroon reflect post-colonial social divisions (Kuchah, 2018), accentuating how LOI choice is highly political and where symbolic power (Bourdieu, 1973) is evident. LOI choice can also raise significant intra-group challenges, especially where one dominant group imposes its values or traditions on others. Prohibiting certain languages in the classroom can also promote intolerance and harmful assimilation policies that can erode individual and group identities (Bourne, 2001). Consequently, development teams must carefully consider the implications that specific LOI choices will have on power and social dynamics in a given context.

It is also important to consider whether the terminology and word choices used in the educational environment foster gender equality and inclusivity. Practitioners must be conscious that the concepts and terminology used in curriculum documents and resources can influence learner cognition, affect and behaviour (Leaper, 2014). For example, different languages present gender in different ways, so curriculum developers and practitioners must reflect on the gender nuances that are conveyed through the language that is used.

In all education settings, not just those relating to EiE, it is vital that a curriculum is developed and delivered in a language appropriate for supporting learner success and inclusion. Across Europe, there are different approaches to this. In some education systems, there is an initial integration phase where language and learning support is provided to newly arrived migrant students in separate classes or lessons. In other jurisdictions, migrant students are placed directly into

mainstream classes but are still provided with additional support (European Commission/EACEA/Eurydice, 2019). However, deciding on a route of classroom-based support is not enough. Curriculum developers must reflect on the language support and limitations that are implicitly and explicitly integrated within the curriculum and assessment approaches. There are many positive effects linked to students' social, cognitive and linguistic development if a curriculum is designed in order to accommodate learning in diverse languages, especially at the primary level (European Commission/EACEA/Eurydice, 2019).

3. Content in locally adapted curricula should be framed so that it is culturally sensitive

During any curriculum development process, developers should be conscious that the content and material referred to in the curriculum is culturally sensitive. Being culturally sensitive refers to a curriculum being relevant, meaningful, respectful and responsive to learners' culture and lived experiences. Although this overlaps with language, this guidance area also focuses on cultural practices, values and histories.

In the case of LP curriculum development, developers should tailor the LP framework through the inclusion of content and examples that are relevant to the learning environment and the learners' individual experiences, needs, interests and worldviews. Gervedink Nijhuis et al. (2013) note that the most well-defined curriculum still falls short if curriculum developers do not critically reflect on whether concepts and activities are culturally sensitive. This is especially important for avoiding clashes between learners' cultural perspectives and more globalised approaches (Deniz & Borgerding, 2018) that are often found in international frameworks.

Across all education contexts, ensuring the curriculum is culturally relevant and that it supports culturally responsive pedagogy is important for rights-based education practices because it affirms students' identities and values in local contexts (Byrd, 2016; Wilson & Alloway, 2013). Addressing students' worldviews and allowing them to engage with local and global perspectives is crucial to creating a positive, inclusive and productive space for learning (Klenowski, 2009). Creating a positive space for learning can also lead to greater learner progress and achievement (Van Laar et al., 2013).

This guidance area can be illustrated through an example related to Mathematics curriculum development. A study in Alaska shows that a curriculum which draws from locally relevant examples relating to harvesting, star navigation, and fish rack construction has a positive correlation with helping students prepare to meet national assessment exams (Kaino, 2013). There is also value in expanding topic areas to include local examples that students can connect to. This helps students to connect an abstract idea to a concrete example. For instance, linking concepts to local plant and animal species, local resources and local environmental sustainability issues can help students more easily grasp complex concepts and models (Hewson, 2012).

The importance of relating science to students' lives is also demonstrated in a study by Albrecht and Upadhyay (2018), who found that local stakeholders believed science is more valuable for their children if it relates to the challenges they may face in their lives. For example, discussions around chemistry and nature are relevant if they help students understand how to respond to the aftermath of natural disasters. Furthermore, a curriculum which helps students respond to their local settings and needs is valuable for both students and families, such as helping students to learn about agriculture-related content to help ensure they have sustainable food sources (Hewson, 2012).

Ensuring cultural appropriateness can be done by modifying or reframing content so that it is respectful, mindful and inclusive. However, this can be a complex task and requires a significant amount of academic, pedagogical and cultural expertise (Atwater et al., 2010). For example, it may be necessary to adapt the framework levelling in order to delay content until an appropriate age according to that culture. However, it should be recognised that modifications may have an impact on coherence across the curriculum leading to additional adaptations being required in order to ensure the quality of learning is upheld.

In some cases, controversial content may be deemed appropriate if framed correctly (Albrecht & Upadhyay, 2018). For example, while reproduction and sex education may be deemed inappropriate and controversial within some cultures, Tripathi and Sekher (2013) have found that in the context of India, teaching sex education with the aim of raising awareness around HIV prevention, sexually transmitted infections and teenage pregnancies has legitimised the importance of introducing this content in formal curricula. It is pivotal that development teams work with and include local experts and stakeholders in the curriculum development process in order to approach these potentially contentious topics effectively.

4. Indigenous knowledge should be included in the locally adapted curricula

The fourth guidance area focuses on developing an LP-based curriculum that includes local indigenous knowledge and indigenous worldviews. However, being aware of and respectful of indigenous knowledge is something all curriculum developers should reflect on. Indigenous knowledge refers to "the understandings, skills and philosophies developed by societies with long histories of interaction with their natural surroundings" (UNESCO, n.d.). For many areas, indigenous knowledge can "inform decision-making about fundamental aspects of day-to-day life" (UNESCO, n.d.). Even if indigenous groups are no longer present in an area, it is important to be aware of and incorporate indigenous legacies and histories within the curriculum in order to provide students with an authentic understanding of how knowledge and culture is shaped and altered.

The process of integrating local indigenous knowledge entails numerous actors and levels. Those involved in adapting and contextualising must also have the experience needed to understand and acknowledge indigenous perspectives (Aikenhead, 2017). Curriculum contextualisation must avoid presumptions, stereotyping, outdated understandings of ontologies and epistemologies, and it must critically engage with values and customs and changes over time (Carey, 2015; Aikenhead, 2017). Identifying and engaging with cultural brokers is also crucial. For example, a study conducted with Syrian refugees in Jordan illustrates the effectiveness of using a Community Readiness Model (CRM) with displaced communities, by portraying its use in evaluating community-based needs through rapid assessment including interviews, focus groups, code mapping, and workshops to identify gaps and understand attitudes (Wells et al., 2019).

Integrating indigenous and local knowledge into a curriculum is crucial for helping learners connect curriculum content to indigenous

practices, decision-making processes, social interactions, rituals and spiritual beliefs. In order to achieve this, curriculum developers must have an awareness of indigenous knowledge, create room for integration and address any conflicts or discrepancies between local views and the dominant academic discourse. While providing access to indigenous knowledge in formal education is important for protecting traditional knowledge, it is also essential for learner engagement. Research in nations facing high immigration and diversity finds that 'gaps' in achievement for learners from minorities may also be linked to the use of curricula which do not affirm learners' diverse identities (Morrison et al., 2008; Wilson & Alloway, 2013). Furthermore, cultural traditions and perceptions at home may appear to be rendered irrelevant if they are not acknowledged in schools, which can lead to cultural degradation (Skutnabb-Kangas & Dunbar, 2010).

It is important for curriculum developers to ensure curriculum content integrates and addresses concepts which are familiar to indigenous perspectives and connecting these to dominant academic discourse. This allows learners to reflect, understand, and negotiate disparities between sources of knowledge (Le Grange, 2007). This is described as cross-cultural pedagogy, where learners engage with both traditional and more globalised academic meanings of a concept side by side (Ng'Asike, 2011). Curriculum developers must also reflect on word choice. Across different languages and cultures, concepts used within the curriculum may be perceived differently. For example, a study conducted by Lee et al. (2012) found that teachers in schools in Taiwan teach time using a cross-cultural approach that is inclusive of Amis indigenous culture.[2] To do this, teachers teach time using approaches familiar to dominant academic discourse (such as time-keeping, clocks, and solar and lunar calendars) and they also present indigenous methods of measuring time in relation to nature or in relation to events and lived experiences. This approach enhances meaning, student awareness and learner engagement (Lee et al., 2012).

Incorporating elements of local indigenous knowledge and culturally sensitive content also helps to deconstruct the hidden curriculum that exists in many learning environments (Wren, 1999). The hidden curriculum can be defined as the unwritten rules, regulations, standards and expectations that form part of the learning process in schools and classrooms, and these elements are not specifically taught to students through the planned curriculum content (Rahman, 2012). In countries such as Australia, Sweden and Canada, studies have shown that the lack of indigenous themes within the prescribed curriculum and the Western values that dominate the hidden curriculum have resulted in lower attainment for indigenous students and a higher level of resistance and withdrawal from formal schooling (Rahman, 2012; Harper & Thompson, 2017; Svonni, 2015). Students come to school with a reservoir of cultural understandings and resources that help them to acclimatise and succeed at school. However, some indigenous students may not possess an awareness of these norms and cultural codes, which can stifle their progress in formal education (Watego, 2005). The process of linking indigenous and Western knowledge systems is effective in engaging indigenous students with mainstream education and in increasing the cultural awareness of everyone involved in the school environment (Rahman, 2012).

2. The Amis tribe is the largest of 14 indigenous tribes in Taiwan recognised by the government. The Amis have their own distinct language, cultural features, traditional customs and social structure (Lee et al., 2012).

5. Locally adapted curricula should support learner well-being, inclusion and success

While the previous guidance areas support learner well-being, inclusion and success, this specific guidance area calls for curriculum developers to explicitly reflect on these areas. Similar to the points above, supporting learner well-being is not just the responsibility of the classroom teacher and school-based support team. Curriculum developers and facilitators must also ensure that the curriculum promotes well-being, inclusion and success for all learners.

Educational spaces, including the curriculum that is taught there, are central to promoting well-being and resilience. For some students, educational spaces also help to restore a sense of normalcy and security for learners and their communities. Curriculum developers must be aware of how curriculum content and expectations positively or negatively impact childhood well-being (CWB). Childhood well-being can encompass a child's developmental progression, including important life events and life transitions (Statham & Chase, 2010). The longitudinal Fragile Families dataset (Fava et al., 2017) provides empirical support for the multidimensional construct of CWB with the following dimensions:

- **Material well-being:** The domain of material well-being may best be described as a measure of financial income, goods, resources, and the ability to provide for basic needs. For education, this relates to ensuring the child has the appropriate resources to engage in the learning environment.
- **Relational well-being:** The relationship domain represents the types of relationships, quality of relationships, and levels of affection expressed towards the child from important people in their lives (e.g., parents, grandparents and close friends).
- **Health and behavioural well-being:** This domain considers the child's physical health, access to health care, and subjective measures of the perception of the child's health by their carers. In addition, more behavioural health aspects are considered (e.g., child feeling sad, lonely, ashamed, and getting into fights with other children).
- **Environmental enrichment:** The environmental enrichment domain focuses attention on aspects pertaining to whether or not a caregiver reads or tells stories to the child, the number of books, toys, puzzles, and instruments that are in the home, and whether or not the child gets to go on outings, or had hobbies.

One dimension not included in the above list, yet important for many young learners and their families, is spiritual well-being (Quosh, 2013; Betancourt & Khan, 2008; Silove, 2013).

When developing a curriculum, whether it is intended for EiE context or otherwise, the above dimensions of well-being should be considered in order to ensure that explicit support for CWB is offered throughout the curriculum and that the curriculum is mindful of local CWB challenges. For example, curriculum developers should reflect on which materials are required for students to progress through the curriculum; what types of relationships are implicitly referred to within the curriculum; and, what assumptions are being made regarding learner behaviours or experiences. Reflecting on these questions can help developers to consider critically how CWB is incorporated and positioned within the developed curriculum. For more details regarding the Learning Passport Social and Emotional Learning (SEL) framework, please see Boyd-MacMillan and DeMarinis (2020).

This guidance area also considers the importance of ensuring access for all students. A core motivation for the development of the LP framework is that all children and young people have the right to good quality education through which they can actively participate and achieve success. However, in many learning contexts that involve conflict and crises, groups of learners are denied access or are not given equal access or support (INEE, 2019; Sæbønes et al., 2015). This includes, but is not limited to, learners with disabilities, cultural groups, language groups and gender groups.

An inclusive approach to curricula involves "a common curriculum for all, based upon differentiated and/or individualised instruction, rather than an alternative curriculum being developed for low achievers" (UNICEF, 2014, p.20). Therefore, curricula should be developed in a manner that ensures the inclusion of all learners including those with disabilities. It is crucial that this inclusive approach is applied to the curriculum development and selection of accompanying teaching and learning approaches, materials and assessments.

Conclusion

Although the Learning Passport project was targeted at supporting displaced learners in EiE contexts, the findings from the project regarding curriculum priorities and curriculum development processes provide a wider contribution. The development of the Framework allowed the development team to focus on the core knowledge needed to support learners across a variety of contexts. The development of the Adaptation Guidance allowed the team to investigate key areas of consideration that should be thoroughly investigated, reflected on and actively incorporated when the framework is developed into a locally based curriculum.

The recommended guidance areas above do not represent a simple, one-time reflective process. Instead, the five guidance areas call for deep and critical reflection that should involve a collaborative process with stakeholders and relevant experts. We recognise that in many curriculum development contexts, this would be a significant shift from the current development process and could lead to potentially difficult conversations and conflicting views. However, the time and resources required to action this adaptation guidance are worthwhile if the aim is to provide a high-quality education provision for all learners. Providing learners with a curriculum that is flexible, supportive and relevant to them will help them to succeed. It will also make the learning environment more inclusive and positive for all learners, which will not only improve learner well-being, but research shows that it will also raise attainment (UNESCO, 2018).

A system-wide approach to providing support is required. This includes, but is not limited to, curriculum development approaches. Beyond curriculum development, the five guidance areas discussed above also serve as valuable areas for reflection when conducting a curriculum and system review. Education ministries should also reflect on education delivery, scheduling, assessment processes, materials, digital tools, as well as professional development provided to curriculum developers, school leaders and teachers to ensure that the education system is inclusive and supportive.

Acknowledgements

We thank Chris Sowton, Dr Eolene Boyd-McMillan, Prof. Valerie DeMarinis, Tim Oates and the Learning Passport External Reference Group for their valuable insights, suggestions and constructive feedback that helped to develop this article.

References

Aikenhead, G. S. (2017). Enhancing school mathematics culturally: A path of reconciliation. Canadian. *Journal of Science, Mathematics and Technology Education, 17*(2), 73–140.

Albrecht, N., & Upadhyay, B. (2018). What do refugee mothers want in urban US school science: learning from the perceptions of science of three Somali mothers. *The Urban Review, 50*(4), 604–629.

Atwater, M. M., Freeman, T. B., Butler, M. B., & Draper-Morris, J. (2010). A case study of science teacher candidates' understandings and actions related to the culturally responsive teaching of 'other' students. *International Journal of Environmental and Science Education, 5*(3), 287–318.

Betancourt, T. S., & Khan, K. T. (2008). The mental health of children affected by armed conflict: Protective processes and pathways to resilience. *International review of psychiatry, 20*(3), 317–328.

Bourdieu, P. (1973). *Cultural Reproduction and Social Reproduction*. Tavistock.

Bourne, J. (2001). Doing 'what comes naturally': How the discourses and routines of teachers' practice constrain opportunities for bilingual support in UK primary schools. *Language and Education, 15*(4), 250–268.

Boyd-MacMillan, E., & DeMarinis, V. (2020). *Learning Passport: Curriculum Framework (IC-ADAPT SEL high level programme design)*. Cambridge University Press & Cambridge Assessment.

Byrd, C. M. (2016). Does culturally relevant teaching work? An examination from student perspectives. *SAGE Open, 6*(3), 2158244016660744.

Cambridge Assessment (2020). *The Learning Passport: Curriculum Framework (Maths, Science, Literacy)*. Cambridge Assessment.

Cambridge Assessment & Cambridge University Press (2020). *The Learning Passport: Research and Recommendations Report*. Cambridge University Press & Cambridge Assessment.

Capstick, T., & Delaney, M. (2016). *Language for resilience: The role of language in enhancing the resilience of Syrian refugees and host communities*. British Council.

Carey, M. (2015). The Limits of Cultural Competence: An Indigenous Studies Perspective. *Higher Education Research & Development, 34*(5), 828–840.

Clark, K. (2017). Are we ready? Examining teachers' experiences supporting the transition of newly-arrived Syrian refugee students to the Canadian elementary classroom. [Unpublished Master's thesis]. *University of Toronto, Toronto, Canada.*

Dearden, J. (2014). *English as a medium of instruction—a growing global phenomenon*. British Council.

Deniz, H., & Borgerding, L. (Eds.). (2018). *Evolution Education Around the Globe*. Springer Berlin Heidelberg.

Elliott, G. (2011). A guide to comparability terminology and methods. *Research Matters: A Cambridge Assessment Publication, 2*, 9–19.

Elliott, G. (2014). Method in our madness? The advantages and limitations of mapping other jurisdictions' educational policy and practice. *Research Matters. A Cambridge Assessment Publication, 17*, 24–28.

European Commission/EACEA/Eurydice, (2019). *Integrating Students from Migrant Backgrounds into Schools in Europe: National Policies and Measures*. Eurydice Report. Publications Office of the European Union.

Fava, N. M., Li, T., Burke, S. L., & Wagner, E. F. (2017). Resilience in the context of fragility: Development of a multidimensional measure of child wellbeing within the Fragile Families dataset. *Children and Youth Services Review, 81*, 358–367.

Forsen, Y., & Guyatt, H. (2015). *Refugee Household Vulnerability Study: Kakuma Refugee Camp*. Kimetrica.

Gervedink Nijhuis, C. J., Pieters, J. M., & Voogt, J. M. (2013). Influence of culture on curriculum development in Ghana: An undervalued factor? *Journal of Curriculum Studies, 45*(2), 225–250.

Greatorex, J., Rushton, N., Coleman, V., Darlington, E., & Elliott, G. (2019). *Towards a method for comparing curricula*. Cambridge Assessment Research Report. Cambridge Assessment.

Harper, A. O., & Thompson, S. (2017). Structural oppressions facing Indigenous students in Canadian education. *Fourth World Journal, 15*(2), 41.

Hewson, M. G. (2012). Traditional healers' views on their indigenous knowledge and the science curriculum. *African Journal of Research in Mathematics, Science and Technology Education, 16*(3), 317–332.

Inter-agency Network for Education in Emergencies (INEE). (2019). INEE Guidance Note on Gender. New York, NY. www.inee.org. Creative Commons Attribution ShareAlike 4.

Jolly, A., Thomas, S., & Stanyer, J. (2020). London's children and young people who are not British citizens: A profile. *Mayor of London Report*. https://www.london.gov.uk/sites/default/files/final_londons_children_and_young_people_who_are_not_british_citizens.pdf [Accessed 17 June, 2020].

Kaino, L. M. (2013). Traditional knowledge in curricula designs: Embracing indigenous mathematics in classroom instruction. *Studies of Tribes and Tribals, 11*(1), 83–88.

Klenowski, V. (2009). Respecting local, cultural contexts for assessment practice in an era of globalization. In E. L. Grigorenko (Ed.) *Multicultural psychoeducational assessment*, 77–94.

Kosonen, K. (2005). Education in local languages: policy and practice in South-East Asia. *First Language First: Community-Based Literacy Programmes for Minority Language Contexts*, 96–132. UNESCO Bangkok.

Kuchah, K. (2018). Early English medium instruction in Francophone Cameroon: The injustice of equal opportunity. *System, 73*, 37–47.

Le Grange, L. (2007). Integrating Western and Indigenous knowledge systems: The basis for effective Science education in South Africa? *International Review of Education, 53*(5–6), 577–591.

Leaper, C. (2014). Gender similarities and differences in language. In T. Holtgraves (Eds.), *Oxford handbook of language and social psychology* (pp.62–81). Oxford University Press.

Lee, H., Yen, C. F., & Aikenhead, G. S. (2012). Indigenous elementary students' science instruction in Taiwan: Indigenous knowledge and western science. *Research in Science Education, 42*(6), 1183–1199.

Marinotti, J. (Ed.), (2016). *Symposium on Language and the Sustainable Development Goals*. Conference proceedings in New York, 21–22 April 2016.

Morrison, K. A., Robbins, H. H., & Rose, D. G. (2008). Operationalizing culturally relevant pedagogy: a synthesis of classroom-based research. *Equity & Excellence in Education, 41*(4), 433–452.

Nerbonne, J., & Hinrichs, E. (2006). *'Linguistic Distances.'* Proceedings of the Workshop on Linguistic Distances, Association for Computational Linguistics, Sydney. 23 July 2006.

Ng'Asike, J. T. (2011). Turkana children's rights to education and indigenous knowledge in science teaching in Kenya. *New Zealand Journal of Teachers' Work, 8*(1), 55–67.

Pinnock, H., & Vijayakumar, G. (2009). *Language and Education: The Missing Link*. CfBT Education Trust and Save the Children.

Quosh, C. (2013). Mental health, forced displacement and recovery: Integrated mental health and psychosocial support for urban refugees in Syria. *Intervention: International Journal of Mental Health, Psychosocial Work & Counselling in Areas of Armed Conflict, 11*(3), 295–320.

Rahman, K. (2012). Belonging and learning to belong in school: the implications of the hidden curriculum for indigenous students. *Discourse: Studies in the Cultural Politics of Education, 34*(5), 660–672.

Romaine, S. (2013). Keeping the promise of the millennium development goals: Why language matters. *Applied Linguistics Review, 4*(1), 1–21.

Sæbønes, A. M., Bieler, R. B., Baboo, N., Banham, L., Singal, N., Howgego, C., McClain-Nhlapo, C. L., Riis-Hansen, T. C., & Dansie, G. A. (2015). *Towards a disability inclusive education*. Background paper for the Oslo Summit on Education for Development, *28*(4), 369–379.

Sandkull, O. (2005, August). *Strengthening inclusive education by applying a rights-based approach to education programming*. In ISEC Conference, Glasgow (pp.1–9).

Setati, M., & Adler, J. (2000). Between languages and discourses: Language practices in primary multilingual mathematics classrooms in South Africa. *Educational Studies in Mathematics, 43*(3), 243–269.

Silove, D. (2013). The ADAPT model: A conceptual framework for mental health and psychosocial programming in post conflict settings. *Intervention: International Journal of Mental Health, Psychosocial Work & Counselling in Areas of Armed Conflict, 11*(3), 237–248.

Skutnabb-Kangas, T., & Dunbar, R. (2010). *Indigenous children's education as linguistic genocide and a crime against humanity?: a global view* (pp.1–128). Kautokeino, Norway: Gáldu.

Statham, J., & Chase, E. (2010). *Childhood Wellbeing: A brief overview* (Briefing Paper No. 1). Childhood Wellbeing Research Centre.

Sugarman, J. (2015). Meeting the education needs of rising numbers of newly arrived migrant students in Europe and the United States. *Migration Policy Institute commentary*. November 2015. https://www.migrationpolicy.org/news/meeting-education-needs-rising-numbers-newly-arrived-migrant-students-europe-and-united-states [Accessed 15 June, 2020].

Svonni, C. (2015). At the margin of educational policy: Sámi/Indigenous Peoples in the Swedish National Curriculum 2011. *Creative Education, 6*(9), 898–906.

Taylor, S., & Sidhu, R. K. (2012). Supporting refugee students in schools: What constitutes inclusive education? *International Journal of Inclusive Education, 16*(1), 39–56.

Tripathi, N., & Sekher, T. V. (2013). Youth in India ready for sex education? Emerging evidence from national surveys. *PLoS ONE, 8*(8), e71584.

UN General Assembly. (1992). *Declaration on the Rights of Persons Belonging to National or Ethnic, Religious and Linguistic Minorities*. 3 February 1992, A/RES/47/135. Accessed 26 January 2018 https://www.refworld.org/docid/3ae6b38d0.html

UNESCO (n.d.) *Natural Sciences—Priority Areas*: http://www.unesco.org/new/en/natural-sciences/priority-areas/links/related-information/what-is-local-and-indigenous-knowledge/

UNESCO. (2018). *Global Education Monitoring Report 2019: Migration, Displacement and Education*—Building Bridges, not Walls. Paris, UNESCO.

UNICEF. (2014). *Conceptualising Inclusive Education and Contextualising it within the UNICEF Mission. Webinar 1—Companion Technical Booklet*. https://www.ded4inclusion.com/uploads/4/7/7/8/47789531/1._contextualizing_ie.pdf

Van Laar, C., Derks, B., & Ellemers, N. (2013). Motivation for education and work in young Muslim women: The importance of value for ingroup domains. *Basic and Applied Social Psychology, 35*(1), 64–74.

Watego, L. (2005). An Indigenous Perspective on Mathematics Contextualisation in a Pre-School: from safety to empowerment. In *MERGA* (Vol. 28, pp.769–774).

Wells, R., Abo-Hilal, M., Steel, Z., Hunt, C., Plested, B., Hassan, M., & Lawsin, C. (2019). Community readiness in the Syrian Refugee community in Jordan: A Rapid Ecological Assessment Tool to build psychosocial service capacity. *American Journal of Orthopsychiatry*. Advance online publication.

Wilson, K., & Alloway, T. (2013). Expecting the unexpected: Engaging diverse young people in conversations around science. *The Australian Educational Researcher, 40*(2), 195–206.

Wren, D. J. (1999). School culture: Exploring the hidden curriculum. *Adolescence, 34*(135), 593–596.

Writing and reviewing assessment questions on-screen: issues and challenges

Victoria Crisp Research Division and **Stuart Shaw** Cambridge Assessment International Education

Introduction

For assessment contexts where both a paper-based test and an on-screen assessment are available as alternatives, it is still common for the paper-based test to be prepared first. The questions are then transferred into an on-screen testing platform to be as similar as possible to their paper-based counterparts. One challenge with this is that, due to the functionality of the on-screen platform, it may not be possible to transfer certain questions; they either have to be adapted or replaced. This creates additional work and potentially raises comparability issues relating to the equivalence of constructs assessed, level of demand and standards. Moreover, the transfer of items into the relevant on-screen testing platform is often undertaken by personnel other than the subject experts who set the questions. In future, if questions are to be used in on-screen tests, it may be advisable to consider their design from the outset, perhaps with questions drafted by setters directly into the preferred testing platform, with these questions then converted into a paper-based test. This could be accomplished first for easier question types. This could reduce or even remove the need to adapt questions later as the questions will already have been designed to work within any restrictions of the on-screen testing platform. There is also the potential that this strategy could minimise risks for comparability as it is likely to mean that the questions could be more similar between modes (though whether all constructs of interest for a particular assessment context can be assessed through an on-screen testing platform would need to be considered).

If such an approach were taken, drafting assessments directly into an on-screen testing platform might require setters and reviewers to have some different skills and knowledge from those required for paper-based assessments, with potential implications for training and recruitment of personnel. In view of this potential approach, this research explored what is involved in setting and reviewing items specifically for on-screen assessments. The research also provides guidance on the training and/or support that setters would need if writing questions into an on-screen testing platform and that reviewers would need if involved in quality-assuring on-screen test materials.

The research brings corroboration and additional detail to previously existing anecdotal insights.

Method

Assessment context and platform

This research was conducted in the context of Cambridge Assessment International Education Checkpoint[1] tests in Mathematics and Science, with a focus on tests for 14-year-olds (Stage 9). As an example of a well-established on-screen testing platform, Inspera[2] was used in the research. It should be noted that the focus of the research was the knowledge, skills and training that assessors would need in order to set or review test questions using an on-screen testing platform; the research is not an evaluation of Inspera, and most of the themes identified in the current research are likely to apply to other on-screen testing platforms. It should also be noted that the data collection was conducted in April 2018 and there have been updates to the software since then.

Participants

Six participants, three with Science expertise and three with Mathematics expertise, were recruited. All had question setting and reviewing experience in their subject area for qualifications such as Checkpoint, GCSE, IGCSE, O Levels and A Levels. The participants represented a range of years of experience in setting and reviewing. None of the participants had previous experience of writing questions into a platform for on-screen testing. Half of the participants had previous experience of reviewing items in an on-screen testing platform.

The participants were assigned identifiers which are used in reporting some of the findings; Sc1 to Sc3 represent the Science participants and Ma1 to Ma3 represent the Mathematics participants.

Procedure

The research was undertaken in two stages.

Stage 1: Advance familiarisation

The aim of stage 1 was to provide participants with some familiarity with the on-screen testing platform and with the relevant curriculum (if they were not already familiar with it). The participants were provided with:

- login details;
- brief notes on how to create a new set of questions within the platform;
- a link to a website with guidance on using the platform;[3]
- the curriculum for Mathematics or Science, as appropriate;
- three or four example questions targeted at the relevant age group, and;

1. Checkpoint tests provide teachers with feedback on learners' strengths and weaknesses to inform their teaching and measure learners' progress over time. They also enable structured reporting to parents. They are available in English, English as a Second Language, Mathematics, Science and Cambridge Global Perspectives.
2. https://www.inspera.com/
3. https://inspera.atlassian.net/wiki/spaces/KB/pages/57311314/Author

- a form in which to record notes on their initial experiences during the familiarisation.

The participants were asked to explore the authoring area of the platform, try entering the test questions provided, make notes in the form on how they found the software and any challenges encountered, and familiarise (or re-familiarise) themselves with the curriculum.

Stage 2: Research meeting

A research meeting was held with each participant individually. This involved a number of activities:

- *Familiarisation with a workload questionnaire*. Participants read some information about the NASA TLX (Task Load Index; Hart & Staveland, 1988). This is a tool used to evaluate an individual's experience of a task in terms of six dimensions of workload: mental workload, physical demand, temporal demand, performance, effort (fatigue), and frustration.
- *Introductory interview*. Participants were asked about their previous experiences of setting and reviewing assessment questions and to describe key points arising from their familiarisation activities in stage 1.
- *Setting activity (approximately 90 minutes)*. Participants drafted some items into the testing platform as if for a Checkpoint Stage 9 test. They were given a printed copy of the curriculum and free choice of topics to use. The researchers observed the question setting from outside the meeting room using Morae remote observation software (TechSmith, 2011). This allowed the researchers to observe unobtrusively the participant's activities on-screen and a small video feed of the participant and desk area.
- *Completing workload questionnaire*. The participants completed the NASA TLX workload questionnaire in relation to the setting activity. This involved indicating the contribution of each workload dimension to their overall workload in both a relative (i.e., by comparing the contribution of pairs of dimensions) and absolute manner (i.e., by rating the contribution of each dimension).
- *Interview regarding setting activity*. Participants were interviewed regarding the setting activity including their process of drafting questions, differences between setting on-screen items and items for paper-based tests, knowledge and skills needed for setting on-screen, and guidance that setters would need to support them in writing good questions for on-screen tests.
- *Reviewing activity (approximately 75 minutes)*. Participants reviewed a set of questions that had been entered into the testing platform in advance. They were asked to review the quality of the items, whether auto-marking worked appropriately, and whether marking guidance was appropriate. Participants were asked to record feedback on a form (electronically or on paper, according to their preference) to report issues with the questions and suggestions for improvement. Again, Morae was used to remotely observe the participants.
- *Completing workload questionnaire*. Participants completed the NASA TLX workload questionnaire in relation to the reviewing activity.
- *Interview regarding reviewing activity*. Participants were interviewed regarding the reviewing activity. The interview covered their process of reviewing questions, differences between reviewing on-screen versus paper-based items, knowledge and skills needed for reviewing on-screen, and guidance that would be needed to support reviewing on-screen questions.
- *Final interview*. Participants were asked whether they felt that the item setting and reviewing that they had conducted during the research would have been similar or not if they had conducted the same work at home.

Results

As described, there were various different types of data collected in this research (i.e., interviews, observations, completed forms, workload questionnaires). The results have been summarised across these types, given that overlapping themes emerged. The sections that follow focus on the general findings from the advance familiarisation (stage 1) and from the research meetings (stage 2). For the research meetings, the findings are organised by setting and reviewing. Note that the completed forms, observations and interviews revealed various (and overlapping) issues and challenges relating to the specifics of using the platform (e.g., how to enter fractions, how to apply a background image to a hotspot item, how to format tables). These have not been reported in full in each section in order to reduce repetition. Instead, a few examples are given where pertinent, and a full list is presented later to indicate points that would need to be included in training or guidance for setters and reviewers working with on-screen assessments.

Advance familiarisation

Based on their reports, most of the participants were able to start entering questions fairly quickly during their familiarisation work. One of the Mathematics participants struggled to get started, however, and sought help from the researchers. All participants experienced some initial challenges with working out how to use certain features of the platform.

A key theme in reports from the Mathematics participants was the issue of being able to set out questions appropriately. The Mathematics participants tended to think that entering questions was more time consuming for Mathematics than it would be for other subjects, presumably due to features such as tables and mathematical notation. Ma2, for example, observed that some features of Microsoft Word that they found helpful were not present in the on-screen testing software. At the end of the exercise, Ma3 felt confident that they could author a routine calculation question but noted that setting out questions might take up more attention than creating the question content when working in an on-screen testing platform:

> *"Currently … I think if I had to author using this system, my time spent would be more on typesetting and less on creating content."*

Ma3 also expressed concern that:

> *"I don't feel I entered any of the given questions completely. There was some kind of issue on all of them. Those questions were designed for pen and paper so you can't just replicate—they've got to be adapted. More creative ways need to be explored, e.g. if in life we are using a pen and paper then it's natural that they'd write down working out. If in life we are using a computer, then you'd use other tools to solve problems, like spreadsheets, or the internet."*

However, this participant was positive about some aspects of the platform's functionality. Whilst exploring the mathematics function in the editing box, Ma3 noted that the software used was TeX,[4] with which they were unfamiliar. After exploring the functionality of TeX, Ma3 concluded that:

> *"TeX has got the power to do quite a bit more than an equation editor."*

The Science participants also experienced some initial challenges with entering questions. One Science participant (Sc3) felt that they worked more quickly as time progressed through the familiarisation activity. However, they reported that they tended to think of a better way to enter a question (e.g., a different item type) after creating it. This suggests that more experience with the platform may aid setters' abilities to write optimal questions. A concern expressed by all participants related to the frustration of knowing the kind of questions they intended to write but being unable to create them:

> *"I often know what I want to do, however, do not seem able to actually do it."* (Sc3)

There were indications that the participants would have liked to be able to ask someone for guidance on the issues they were experiencing:

> *"At the end, thought I had grasped some of the basics but was left a little frustrated that I could not sort out some relatively minor issues on my own—really needed someone to ask."* (Sc3)

Experiences with the platform during individual research meetings: setting

There were some commonalities in the process setters used when setting questions in the platform as part of the activity, but also some variations:

- Three participants (Ma1, Sc2 & Sc3) started by viewing the curriculum document and either selecting one or more curriculum points to assess or coming up with a question idea at this point. Broadly, this led into consideration of how the question idea or the curriculum point could be assessed using the available item types. Sometimes, the items were then mapped out in notes on paper either with draft wording or just as an idea or rough structure for some parts of a question. Then, participants began inserting items in the platform and drafting question content. Two of these setters usually completed a draft of an item before continuing to the next item, but one setter (Sc2) inserted several items in quick succession and then moved between them adding and editing content.
- Two participants (Ma2 & Ma3) began by selecting an item type to try out. They then viewed the curriculum and selected a curriculum point that they could assess with this item type. One of them then paused whilst they mentally planned the question (Ma3), which was reported to replace their usual and more natural process of developing ideas as they began to write in Word. Participants then inserted and drafted items in the platform. Participants felt that starting by selecting an item type was not an ideal strategy for writing questions but that they were led to this strategy by setting within the platform.
- The remaining participant (Sc1) used a mixture of the two broad strategies described—sometimes beginning by viewing the curriculum but sometimes starting by selecting an item type. Their approach was quite exploratory in terms of investigating different item types and abandoning them if they were not appropriate.

Each participant's process was mapped out as a flow diagram of the steps that they reportedly took to create items during the activity. These are shown in Figure 1. (Note that these do not represent the number of items that participants created.)

Participants experienced a range of challenges during setting. In particular, there were issues relating to the use of 'matching' and 'drag and drop' task types such as locating the tools needed to control features of the questions. Some participants experienced issues with: how to set up auto-marking to give appropriate marks, copying and pasting images, setting up the layout so that there is an initial stimulus before the items, and using TeX. Science participants experienced some problems with placing units after an answer space and with using subscript and superscript in response options (e.g., chemical formulae in multiple choice questions).

Despite some challenges, participants explored and experimented within the platform, for example, trying out different item types, and creating innovative questions (e.g., testing knowledge of the number of electrons in each shell for a particular element using the 'matching' item type). There were positive indications that participants hoped to get better at using the platform with more time. For example, one participant reported that their confidence increased as the session went on.

Differences compared to writing for paper-based tests

During the interviews, participants were asked how setting on-screen items compared to writing for a paper-based test. Participants reported that setting items in the platform made them feel more restricted as their lack of familiarity with the software acted as a barrier. There appeared to be some avoidance of attempting to draft test questions that they were not confident in how to create. Some participants felt that the item types affected their choices around the questions to write. One participant reported writing shorter questions than normal and that they did not develop questions by adding more parts in the same way that they would normally. There seemed to be a tendency to use as many auto-marked questions as possible in order to make use of the advantages of on-screen tests, even though this might change the kinds of questions set compared to their usual setting. It was noted by some that dealing with the software took up part of their thinking and reduced their focus on the subject.

There was some concern about curriculum coverage in that it might be difficult to test some topics or skills within the platform, or that it might simply be more time consuming to set questions on some topics or skills, which would put setters off creating them (especially if payment is per question or per mark rather than per hour). It was also suggested that setters may be less likely to try innovative questions because there is more chance of not being able to create them successfully.

Workload experienced during setting

Participant frustration during setting was a recurring theme across the data collected, a concern that was particularly manifest in the outcomes of the workload questionnaire. As already described, after the setting

4. TeX is a formatting system that is useful for typesetting mathematical formulae, see https://en.wikipedia.org/wiki/TeX (retrieved 26 June 2020). It is provided within Inspera for this purpose.

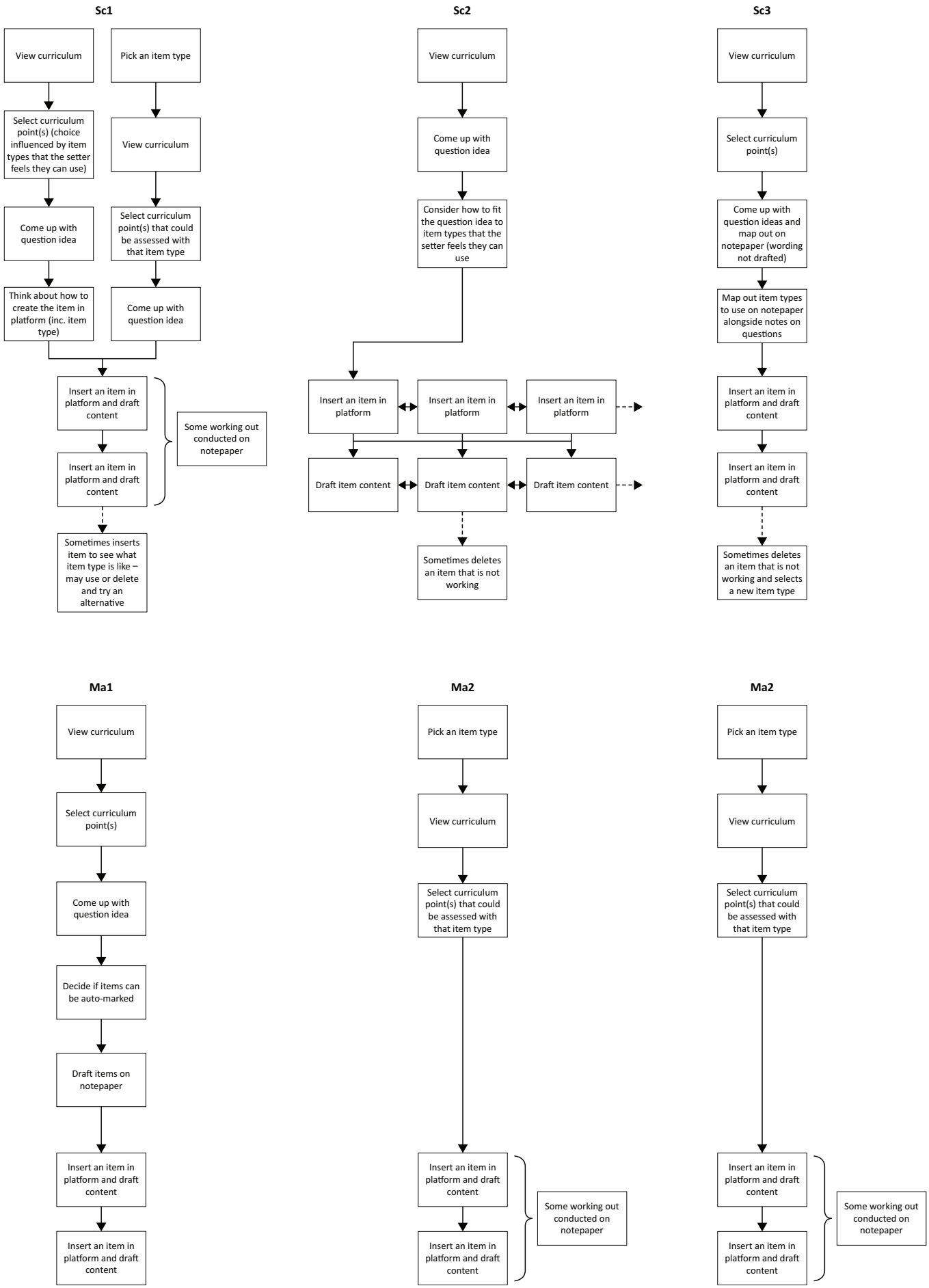


Figure 1: Flow diagrams showing processes during the setting activity.

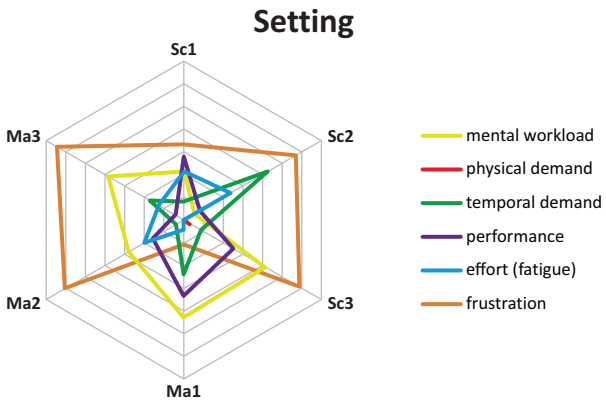


Figure 2: Profiles showing the workload reportedly experienced by participants during setting.[5]

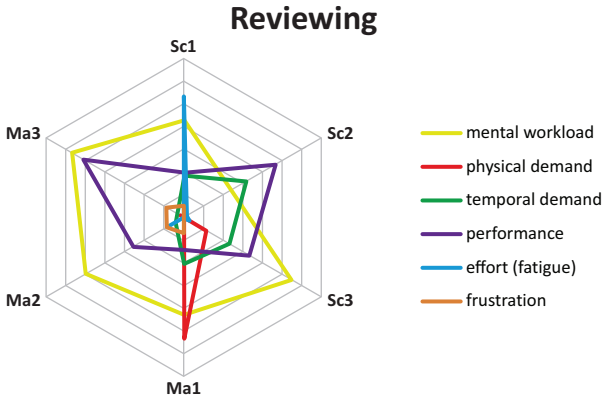


Figure 3: Profiles showing the workload reportedly experienced by participants during reviewing.[5]

activity (and the reviewing activity) participants were asked to report on the workload they experienced in relation to six workload dimensions in both an absolute and relative way.

Workload was analysed according to the usual procedure for NASA TLX workload responses, which for each individual produces workload values for each dimension for setting and for each dimension for reviewing. The values give a measure of the perception of each dimension at that time relative to the other dimensions. There is an underpinning assumption that there is a maximum possible workload and that a high value on one dimension will mean a lower value on one or more other dimensions. In effect, one dimension can 'crowd out' others. It should be noted that as the measures are based on personalised reflections, it is not appropriate to directly compare between individuals. However, common patterns can be considered and comparisons made between activities for one individual.

Figure 2 shows the profiles of the workload experiences reported by the participants during setting. Points further from the centre represent a higher experienced demand for that individual participant relative to the other dimensions. Figure 2 illustrates that for most participants, frustration was the most prominent part of the workload experienced during setting. This workload dimension is defined in terms of how insecure, discouraged, irritated and annoyed the individual feels rather than secure, gratified, content and complacent. These measures are consistent with participants' comments around feeling limited in what they could do due to lack of familiarity with the platform. Mental demand (the amount of mental and/or perceptual activity that is required, for example, thinking, deciding, calculating, remembering, looking, searching) appears to be the second most strongly experienced demand looking at the group of participants as a whole, and was the most strongly experienced workload dimension for one participant. This makes sense given the nature of what is involved in setting test questions.

For one participant, temporal demand was also quite high. This relates to the amount of pressure felt due to the rate at which the task elements occur and to feelings of being rushed. This could suggest that they felt some pressure due to the research situation. With regard to setting, it may be that experiences of frustration have potentially reduced or crowded out other factors such as mental workload and performance for most participants.

5. The axes of Figures 2 and 3 go from 0 (at the centre) to a value of 35 (outer hexagon).

Experiences with the platform during individual research meetings: reviewing

The observations, participant feedback notes about the questions, and interviews provided various insights into the process of reviewing on-screen items. Participants viewed each question checking for various issues, tried out various possible responses, and (usually) viewed the marking guidance. Some participants reported using their usual review process of reading through all the questions first and then conducting a second pass to consider the marking guidance.

When asked about their process of reviewing items, most participants' comments focused on the aspects of items that they were checking. All had a number of elements in mind which appeared to be acting as a checklist as they worked. These included the match to the curriculum, the accuracy of the subject content, spelling and punctuation, appropriateness of command words, formatting and layout, and the functioning of any auto-marking. Some participants considered alternative options for how the questions could be structured, how the stimulus relates to the different parts, and how the stimulus and related items should be set out.

Differences compared to reviewing for paper-based tests

When asked how reviewing questions on-screen compared to reviewing items for a paper-based test, participants generally felt that the process was quite similar in that they were looking for similar things. Some participants reported that they usually reviewed Word documents of paper-based tests and their mark schemes on screen, making the process more similar than if they worked on printed copies. Some differences were mentioned such as considering whether the item type was appropriate, trying out answers for auto-marked items and moving between views in order to check how the item would appear and to check the marking guidance. It was noted that: "What I'm reviewing is different but...what you're trying to do is pretty much the same." One participant felt that reviewing on screen encourages more focus on the marking criteria than on the question.

Workload experienced during reviewing

Figure 3 shows that for reviewing, participants tended to experience a high mental workload (e.g., thinking, deciding, looking) and a fairly high focus on performance (how successful the individual thinks they are in what they have been asked to do and how satisfied they feel with what they accomplish). In other words, participants tended to experience the task as quite challenging in terms of the mental or perceptual activity

involved, and they felt a fair degree of satisfaction in their performance. For one participant, the fatigue of carrying out the reviewing activity was the most strongly contributing factor to their workload, and for another participant, physical demand was most prominent. The reasons for these spikes are unknown. Comparing the two graphs in Figures 2 and 3 (whilst keeping in mind how the values represent relative experiences of the workload dimensions), the clearest difference is that experiences of frustration were generally higher for setting than reviewing. This appears to be offset by greater awareness of mental workload demands and better perceptions of performance during reviewing.

Issues affecting both setting and reviewing

A number of issues affected both setting and reviewing.

Mark allocation

When asked in relation to setting, participants reported that using the platform did not change the mark allocation that they felt was appropriate but that sometimes they did not know how to apply auto-marking so that the mark allocation was correct. Sometimes this led them to adjust the question so that it was worth a mark allocation that they could set up.

Similarly, in relation to reviewing, participants felt that using the platform did not change the mark allocation that was appropriate. Some noted that the software appeared to have affected the mark allocation of some of the items they reviewed, and that the platform might affect what is possible in terms of mark allocation. An interesting comment from one participant was that where a test will be available in both modes, if an auto-marked item gave students more support (e.g., provided response options when the paper-based equivalent item did not) then there could be a case for the on-screen item being worth fewer marks than its paper-based counterpart.

Auto-marking

Participants were asked whether they experienced any challenges relating to auto-marking. Setting up auto-marking was said to be unproblematic for straightforward item types.

One participant reported initial difficulty locating where to put the correct response for auto-marked questions. Another felt that setting up auto-marking was simpler than creating a mark scheme. Participants commented that the need to enter every correct and incorrect spelling that could be accepted was an additional challenge compared to usual. Participants liked that they could try out the auto-marking.

In relation to setting items, participants expressed a need for familiarity with what each item type can facilitate in terms of marking. Specifically, participants mentioned issues with setting up marking for drag and drop items. One participant mentioned that there were challenges with setting up auto-marking where some responses could appear in more than one order (e.g., the products of a chemical reaction in a chemical formula question). Such issues could mean that some questions with short answers might still need to be set up as manually marked questions.

In terms of reviewing auto-marking, most participants felt that checking auto-marking was not problematic and that it was easy to try different responses and see how they were marked. The only challenge mentioned was the need to try out a variety of different options and alternative wordings (e.g., is 'the rope' accepted as a correct answer as well as 'rope'), echoing comments in relation to setting.

Marking guidance

Participants felt that writing marking guidance for manually marked on-screen items would be very similar to writing a mark scheme for a paper-based question. One participant questioned whether a three-column layout, similar to their usual mark schemes, would be needed and how to achieve this (e.g., inserting a table).

When reviewing, some participants initially had difficulty locating the marking guidance but they tended to get better at this with practice. However, they sometimes forgot to check the marking guidance for manually marked questions. Some considered it a challenge or tiring to move between areas. One participant felt that it would be easier to have a mark scheme for the test as a separate document.

Pre-requisite knowledge and skills needed by setters and reviewers

Participant comments suggested that there were very few specific knowledge and skills that were pre-requisites to setting or reviewing on-screen items, beyond the knowledge and skills needed to set or review items for paper-based tests. The pre-requisites that participants mentioned related to general computer skills and confidence, patience and imagination (in order to be able to write questions in a different way).

Training and support materials and what they should cover

In terms of necessary knowledge and skills that could be learnt for the purpose of the role, becoming familiar with various aspects of the platform and its functionality was the main theme. Most participants expressed a preference for face-to-face training, perhaps including some guided learning followed by trying out setting questions with support when needed. One participant felt that some initial independent familiarisation followed by an opportunity to discuss queries would be sufficient. Training could be shorter for reviewers who need to know what the system is capable of but do not need to be able to enter items.

Guidance documents were also mentioned by most participants as potentially helpful with some comments that these should be structured and signposted such that users can easily navigate to find points that they need. There was an interest in having access to some example questions already entered into the platform, and their paper-based counterparts, as a way of seeing how questions can be adapted.

Given the challenges experienced by the participants during the research, setters appear to need training and/or guidance on the following:

- **Setting up the question**
  - The terminology used in the relevant platform;
  - The platform's item types and how to use them (e.g., how to enter fractions or formulae as labels in matching items);
  - How to select an appropriate item type for the question to be asked, especially for common questions (e.g., completing a word equation in chemistry, labelling a diagram);
  - How to test skills relating to handling data and understanding geometrical figures;
  - How to control the layout of questions (e.g., how to display a stimulus alongside the items);

  - How to enter mathematical content (e.g., fractions, formulae);
  - How to create and insert diagrams/graphs/other visuals;
  - How to format tables (e.g., bold in headings, text alignment, width of columns);
  - How to place a unit at the end of the answer line, or text at the start of an answer line (e.g., 'x = _____');
  - How to allow students to enter subscript or superscript into answer spaces (or awareness that this is only possible in certain item types);
  - Which command words should be used (e.g., 'type' or 'write');
  - Whether setters should address formatting or whether this will be checked and refined by a typesetter;
  - How to view stimulus material or earlier items alongside a new item when setting;
  - Whether randomisation of the order of response options should be left on or turned off;
  - Whether 'true'/'false' questions can be used in the test.

- **Setting up marking**
  - Whether a certain proportion of items/marks should be auto-marked;
  - Which kinds of questions common in a subject (e.g., rounding questions in Mathematics, Mathematics questions where the response is a recurring decimal) can be auto-marked and which should be manually marked;
  - How to set up auto-marking to award the appropriate number of marks (e.g., for a matching question where two marks should be given for four correct matches);
  - How to identify all possible correct answers and incorrect spellings that can be accepted (for auto-marked items);
  - Location of marking functions in the platform (e.g., where to set up auto-marking or where to enter marking guidance);
  - How marking guidance should be set out (e.g., in columns);
  - Whether marking guidance should be entered for auto-marked questions.

As mentioned, training for reviewers may not need to be as detailed but some understanding of the platform is needed. Based on insights from the research, training or guidance for reviewers should include:

- the platform's item types and how they can be used;
- navigation in the platform, in particular, locating marking guidance, locating question previews;
- how best to set out screens or move between screens during reviewing such that both auto-marking and marking guidance are checked (e.g., using two internet browser windows so that the item and marking guidance can be viewed side-by-side);
- whether the reviser is responsible for feeding back on formatting issues relating to how the item appears on-screen or whether this will be checked by a typesetter;
- how to identify all possible correct answers and incorrect spellings that can be accepted (for auto-marked items);
- what is possible in terms of awarding marks in auto-marked items.

Discussion

Whilst this research was conducted with one specific on-screen testing software platform and some findings may be specific to that platform, many insights seem likely to reflect issues that setters and reviewers would encounter when beginning to use any on-screen testing platform.

A number of key issues arise from the current research. Firstly, participants varied in how quickly and easily they appeared to be able to get started with entering questions into the platform, suggesting that some setters and reviewers would need more familiarisation and support than others to use an on-screen testing platform.

Participants reported feeling somewhat restricted when setting items due to their lack of familiarity with the platform and that dealing with the software reduced their focus on the subject to be assessed. This aligns with the frustration that appeared to be salient in responses to the workload questionnaire. These issues could reduce with time and increased familiarity with the platform, but possible effects on the questions produced should be considered. There was some evidence that setters may avoid certain item types (potentially reducing curriculum coverage or coverage of assessment objectives), write shorter questions than normal, write less creative questions and, potentially, produce lower quality questions. Any such limitations on the questions produced could mean that a test would not cover the full range of constructs that it was intended to assess (which could in turn affect teaching). The setting process may also be slower, particularly for more complex item types, or if the setter needs to change the item type part way through drafting a question. These points would need careful management in the early days of setters being asked to draft questions into an on-screen testing platform, to ensure the production of a good bank of items assessing the full range of relevant constructs.

Whilst one of the aims of moving towards asking setters to set into an on-screen platform would be to remove the need for questions to be adapted later for on-screen testing, the possible effects of this on the setter's process need to be considered. In the current research, some setters began their process of setting by selecting an item type to use and then working out what content in the curriculum they could test. Whilst research into question setting for paper-based tests indicates that the type of question to be used is often considered alongside the content to be tested when beginning to plan a question (Johnson et al., 2017), the current participants did not feel it was ideal for question setting to be driven by item type. The strategy of starting with an item type might reduce as setters become more familiar with the platform, but if this strategy were to continue to be used at least some of the time, the consequences would need to be considered.

Overall, it seems that it would be possible for setters to create at least some of their questions within an on-screen testing platform such as Inspera. Setting in this way has the potential advantage of avoiding the need for questions that were written for a paper-based test to be substantially versioned or even replaced for an alternative on-screen test, thus saving time and resources, whilst also minimising risks relating to comparability. However, training and guidance would be needed to support setters, and the time and cost investment in this needs to be considered. In addition, there may be initial frustration involved in setting into an on-screen testing platform since it may make setting slower and it may limit the nature of the questions that are produced, especially in the early days. Setter satisfaction should be taken into account as frustration could be an issue.

Another issue to consider is whether it is likely to be productive to ask setters to create all types of questions into the on-screen platform. There is a potential risk, for example, of setters tending to select only item types that are easy to use. Relatedly, whether to ask setters to create questions that would require additional training beyond the basic platform should be considered. For example, within Inspera, an application called GeoGebra can be used to create more sophisticated questions involving graphics, where the aim is to ask students to draw something or add to a diagram. Using GeoGebra would require additional training and it might or might not be worth the setter undertaking this training when balanced against the frequency with which they would use this functionality. An alternative would be to give setters an awareness of what GeoGebra can do, and give them the option to draft some questions in word-processing software with the question later created in the platform by a typesetter with GeoGebra training.

Returning to the finding that participants felt restricted by the platform, it was apparent that sometimes they had ideas for questions that they found they could not implement within the platform. Whilst participants sometimes explored innovative ways to assess concepts, sometimes the restriction they experienced led to compromised decisions about question design that were not satisfactory to participants. This could suggest there is potential for a situation where it is not possible to create questions that tap into certain parts of learning. Over time, if setters can no longer create certain kinds of questions that they would usually write, this could adversely affect content coverage and construct representation. If some individuals are unwilling to make such compromises, they may drop out of involvement in setting. New setters would then be recruited, who might be more accepting of the compromises, thus perpetuating a gradual change in the constructs being assessed. Care would be needed to mitigate risks of this kind in terms of ensuring comparability over time and representation of the constructs contained in the curriculum or syllabus. Asking setters to record question ideas that they could not implement and then working with the software developers to implement appropriate revisions would be one possible way forward.

References

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp.239–250). North Holland Press.

Johnson, M., Constantinou, F., & Crisp, V. (2017). How do question writers compose examination questions? Question writing as a socio-cognitive process. *British Educational Research Journal, 43*(4), 700–719.

TechSmith. (2011). *Usability testing for software and websites*. TechSmith Corp.

A way of using taxonomies to demonstrate that applied qualifications and curricula cover multiple domains of knowledge

Irenka Suto, Jackie Greatorex, Sylvia Vitello Research Division and **Simon Child** Cambridge Assessment Network Division

Introduction

Although they can sound rather grand and esoteric, educational taxonomies are essentially schemes of classification. They are often hierarchical, and provide the terminology that educationalists need to describe and work with different areas of knowledge (so-called 'knowledge domains'). Ever since Bloom and his colleagues created their seminal taxonomy of educational objectives (Bloom et al., 1956; Krathwohl et al., 1964), it has widely been considered good practice to use taxonomies to formulate and review curricula, learning objectives, and associated assessments. Demonstrating sufficient coverage of each of an adequate range of knowledge domains and subdomains is critical for authenticity, for assessment reliability, and for transparency surrounding what students are learning. It is important for regulators, employers and university admissions tutors, as well as the students themselves, to recognise the areas of knowledge, skills and understanding that have been taught and mastered in a particular course.

Bloom et al. (1956) initially created a taxonomy which focused on the cognitive domain. That is, it classified thinking skills as relating to knowledge, comprehension, application, analysis, synthesis, and evaluation. Since Bloom's original work, revisions have been published and alternative taxonomies have been developed to accommodate advances in psychological understanding. Bloom et al. (1956) originally intended to go beyond the cognitive domain, creating a psychomotor taxonomy that focused on physical development. Although they never realised this ambition, some of the more recent taxonomies have done so, covering multiple domains or different single domains. For example, another non-cognitive domain which has been included in some taxonomies is that of interpersonal knowledge, skills and understanding (Hutchins et al., 2013).

Broader domain coverage is important given that many professions and career paths draw upon several different types of knowledge (Bandaranaike & Willison, 2015; Laxmisan et al., 2007; Savic & Kashef, 2013). It seems probable that *general* (sometimes called *academic*) educational taxonomies that cover both cognitive and non-cognitive knowledge domains may also be applicable in applied educational contexts. However, this wider applicability of such taxonomies is relatively underexplored.

In many countries, applied curricula and assessments (often described as 'vocationally related' curricula in England) are perceived as the 'poor relations' to their more academic equivalents (Gleeson & O'Flaherty, 2013; Kämäräinen & Fischer, 2008; McGrath et al., 2006). When academic staff select new students for places in universities and other higher education institutions, they may be unaware of the cognitive demands of applied qualifications. Clarifying such cognitive demands could facilitate the progression of applied students to higher education. Similarly, clarifying the non-cognitive demands could facilitate progression into employment and vocational training schemes. Moreover, identifying shared domain coverage between academic and applied curricula could help to bridge the gap in esteem that is often found between general and applied education routes.

The broad aim of the present study was to explore whether any educational taxonomies that were designed for general educational contexts could be utilised in applied educational contexts. Below, we describe how we identified published taxonomies with sufficient potential and selected the most appropriate. This process led us to develop a new model of demand. We then applied the selected taxonomies experimentally to existing curricula in a range of applied subjects which are taught at secondary and tertiary level in England. We also used the selected taxonomies to develop a tool for writing educational objectives. Finally, we offer some suggestions for applying this approach in other areas of assessment.

Selection of taxonomies

Through a review of the literature, nine published taxonomies were initially identified as having sufficient potential to be utilised in applied contexts. The first four cover multiple domains:

1. Anderson et al. (2001) and Krahwohl (2002)

This taxonomy is grounded in cognitive psychology and is a revision of Bloom's taxonomy. One major difference, however, is that Bloom's taxonomy has one dimension whereas this revised version has two dimensions. The first dimension comprises levels of cognitive processing, ranging from low to high complexity, namely: remember; understand; apply; analyse; evaluate; and create. The second dimension comprises levels of knowledge, ranging from concrete to abstract, specifically: factual; conceptual; procedural; and metacognitive.

2. Atkinson (2013)

Atkinson (2013) adapted several taxonomies to form a more comprehensive framework. He drew together taxonomies of affective, cognitive, psychomotor and knowledge domains (Anderson et al., 2001; Dave, 1967; Krathwohl et al., 1999). Each domain is hierarchical in the sense that students generally achieve a lower category within the taxonomy before they achieve a higher category. For example, students must be able to comprehend factual information before they can apply it to real world contexts and problems.

3. Hauenstein (1998)

Hauenstein (1998) provides a holistic taxonomy, which is a combination of the affective, cognitive and psychomotor domains. The taxonomy categories are ordered in terms of learning, development and complexity. From lowest to highest they are: acquisition (gaining new knowledge); assimilation (integrating new knowledge with what is already known); adaption (adapting knowledge to solve problems); performance (analysing, qualifying and evaluating information and situations); and aspiration (synthesising, hypothesising, resolving complex problems and striving to achieve higher levels of expertise).

4. Marzano and Kendall (2007, 2008)

Marzano and Kendall (2007, 2008) present a taxonomy comprising two dimensions: knowledge domains and mental processing.[1] Within the first dimension there are three different domains: information (declarative knowledge); mental procedures (procedural knowledge); and psychomotor procedures. There is no hierarchical relationship amongst these domains. Within the second dimension there are several levels of mental processing, which are hierarchical. From the lowest to the highest, these are: retrieval; comprehension; analysis; knowledge utilization; metacognition; and self-system (beliefs and motivations determining the level of engagement).

The remaining five taxonomies each focus specifically on a single domain:

5. Carpenter and Wisecarver (2004)

Carpenter and Wisecarver (2004) offer a taxonomy of interpersonal performance in the workplace. They used the literature to propose an initial taxonomy, tested it empirically, and updated it accordingly. The resulting taxonomy includes knowledge and skills related to rewarding, influencing, managing, and formal staffing. The interpersonal domain is unusual in that the different categories within it are not cumulative; that is, it is not necessary to have mastered one category in order to master another.

6. Hutchins et al. (2013)

Hutchins et al. (2013) used the literature about taxonomies and training to construct a comprehensive taxonomy of interpersonal skills. The four high-level skill groupings in the taxonomy are: interpersonal communication skills; relationship building skills; peer leadership skills; and social/behavioural agility skills. There is no hierarchy amongst these four groupings, and each comprises further subcategories of skills.

7. Klein et al. (2006)

Klein et al. (2006) provide a taxonomy of interpersonal skills (communication and relationship building). Cognitive theory underpins their thesis that several factors contribute to the perceptual and cognitive processing that produces interpersonal performance. These factors are: life experience; individual differences; motivation; the environment (such as roles and local rules); and plans.

8. Harrow (1972)

Harrow (1972) developed a taxonomy for the psychomotor domain based on theories of children's psychomotor development. The classifications (from lowest to highest level) are: reflex movements (responses to stimuli without conscious cognition, for example, postural adjustment); basic-fundamental movements (combinations of reflex movements, for example, bending); perceptual abilities (interpretation of stimuli, which is used to adjust the environment, for example, dodging a

1. What these authors mean by 'mental processing' has been termed 'cognitive processing' or 'intellectual processing' by other authors, for example, Atkinson (2013).

moving ball); physical abilities (activities requiring sustained strenuous effort/muscular extension/wide range of movement at the hip/quick precise movements, for example, wrestling); skilled movements (efficient complex movements, which manipulate basic-fundamental movements, for example, sport/dance); and non-discursive communication (communication through body movements, for example, facial expression/dance movements).

9. Biggs and Collis (1982)

Biggs and Collis (1982) developed the Structure of the Observed Learning Outcome (SOLO) taxonomy. The SOLO taxonomy is based upon Piaget's levels of child development, from concrete to abstract cognitive processing. The categories within the taxonomy are: pre-structural (responses miss the point and the approach is too simple); uni-structural/multi-structural (aspects of the assessment task are completed); relational (the response to the task is an integrated whole and shows a satisfactory understanding of the topic); and extended abstract (the understanding of a topic is abstracted and can be generalised to a new topic area).

In order to justify our ultimate choice of taxonomy for use in applied educational contexts, we reviewed each of the above nine taxonomies using six pre-determined inclusion criteria:

1. Credible in terms of its underpinning theory and/or empirical basis.
2. Broad enough to incorporate a sufficient range of knowledge domains:
   (i) information /declarative knowledge;
   (ii) mental procedures;
   (iii) psychomotor procedures; and
   (iv) interpersonal knowledge.
3. Hierarchical or cumulative, such that higher levels tend to relate to higher grades in assessments.
4. Straightforward enough for routine use by assessment developers with little first-hand research experience.
5. Written accessibly.
6. Readily available.

None of the nine taxonomies was found to meet all the selection criteria. This is primarily because no individual taxonomy incorporated a sufficient breadth of domains. Taxonomy 4 by Marzano and Kendall (2007) came closest, covering the information domain ('declarative knowledge'), the mental procedures domain, and psychomotor procedures domain. It was selected for use in conjunction with Taxonomy 6 by Hutchins et al. (2013), which provided the most comprehensive articulation of the interpersonal domain. The standard application format of Marzano and Kendall (2007) was found to be readily extendable to Hutchins et al. (2013).

Knowledge domains covered by the selected taxonomies

Table 1 summarises the four domains of knowledge covered by the two selected taxonomies. As mentioned previously, it is important to note that these four domains cannot be described as hierarchical relative to one another. For example, the psychomotor domain cannot be said to be either more or less demanding than the information domain. All individuals will vary in terms of the domains in which their strengths and weaknesses lie.

Table 1 also indicates that the four domains can be subdivided into categories of knowledge (Marzano & Kendall, 2008). There are five categories of information, four categories of mental procedure, three categories of psychomotor knowledge, and four categories of interpersonal knowledge. In the first three domains, these categories are hierarchical and cumulative in nature, whereas in the interpersonal knowledge domain, they are non-hierarchical.

It is also important to note that the knowledge domains (and the categories of knowledge within them) do not have a homogenous uncompounded nature. Instead they comprise many subdomains of knowledge, which relate to different subject disciplines. Within the psychomotor domain, for example, the psychomotor skills and understanding of an expert violinist differ from those of an expert antique furniture restorer. Both types of expertise result from many hours of education and experience, but the skills and understanding entailed are not interchangeable. Violinists cannot automatically restore antique furniture, and vice versa. It is possible that students with an aptitude for a subject that draws extensively on a particular subdomain will also have an aptitude for subjects that draw upon other subdomains within the same domain. Some students are generally 'sporty' whilst others are generally 'good with people' for example. Other students may have an aptitude for learning foreign languages. In general education and assessment, subjects that frequently go together in this way and draw heavily upon similar subdomains of knowledge are often known as 'cognate' subjects.

Levels of mental processing covered by the selected taxonomies

Syllabuses and curricula are often articulated in terms of learning objectives. Concurring with Bloom et al. (1956) and other authors, Marzano and Kendall (2008) argue that a learning objective should make reference not only to a specific domain (or domains) of knowledge, but also to the student behaviours that would provide evidence of the level of understanding or skill relating to that knowledge domain. These behaviours are displayed in a student's performance during assessment, and reflect the sophistication of the student's internal mental processing. The higher the level of mental processing required in an assessment task (i.e., the more complex the performance requirements), the greater the demand placed on the student.

Marzano and Kendall (2007, 2008) propose six levels of mental processing (Table 2). We found that these can be applied to all four of the knowledge domains discussed previously, including the interpersonal domain articulated by Hutchins et al. (2013). As shown in Table 2, the six levels form a hierarchy of demand.

All six levels are subdivided into 'operations', which are arranged hierarchically and cumulatively. That is, lower operations are encompassed by higher operations. The lowest level of mental processing, retrieval, is about turning our attention to that which we know about but are currently not thinking about (Marzano & Kendall, 2008). As Table 2 shows, retrieval is divided into three operations: recognising, recalling, and executing (Marzano & Kendall, 2008). An example of the cumulative relationship amongst the operations is that of cake-making. Students must be able to decide whether a recipe is accurate, and recall details of the methods stated in the recipe, before they can bake a cake.

The second-lowest level of mental processing is comprehension (Marzano & Kendall, 2008). It is about identifying the key or defining characteristics of knowledge. As indicated in Table 2, there are two

Table 1: Summary of the four knowledge domains and their key features.

Domain	*Domain description*	*Categories*	*Category definition*
Information (Marzano & Kendall, 2007, 2008)	Declarative knowledge. Vocabulary. Factual knowledge such as technical vocabulary. The 'what' of human knowledge.	Principles	Specific types of generalisations focusing on cause–effect or correlational relationships.
		Generalisations	Statements for which examples can be given.
		Time sequences	Include key events that happened between two points in time.
		Facts	Give information about people, places, things and events.
		Vocabulary terms	Phrases learners understand accurately.
Mental procedures (Marzano & Kendall, 2007, 2008)	Mental procedures detailing how to do something: in situation X follow action Y. The 'how-to' of human knowledge.	Macro-procedure	Highly robust mental processes that involve the execution of many interrelated subprocedures.
		Tactics	A set of several mental general rules with a general pattern for the order in which the rules are executed.
		Algorithm	Mental procedures comprised of specific steps which are consistently and automatically applied.
		Single rule	Such as *'IF-THEN'* (Marzano & Kendall, 2008, p.13).
Psychomotor procedures (Marzano & Kendall, 2007, 2008)	Physical procedures, such as being able to serve in tennis.	Complex combination rules	Groups of simple combination procedures interacting and happening simultaneously.
		Simple combination rules	Groups of foundation procedures interacting and happening simultaneously.
		Foundational procedures	The ability to use your body.
Interpersonal knowledge/ skills (Hutchins et al., 2013)	Knowledge and skills used when people are interacting with one another.	Interpersonal communication skills	Express and assimilate information in social interaction. This involves listening, speaking, writing, sending/receiving non-verbal signals in an empathetic, attentive, responsive and confident manner.
		Relationship building skills	Develop & keep relationships with others, to support others, & build strong beneficial alliances as well as manage & resolve conflicts.
		Peer leadership skills	Coaching, counselling, motivating & empowering group members. Gladly interact with a team, earn trust & respect, dynamically participate in problem solving & decision making.
		Social/ behavioural agility skills	Monitor & interpret our own and other's behaviours & modify self-presentation during social interaction to influence & control the interaction.

operations that comprise comprehension: integrating and symbolising. As with retrieval, the operations are cumulative (Marzano & Kendall, 2008). For instance, healthcare students must be able to reduce and organise information about certain pharmaceutical drugs down to key characteristics before they can represent the knowledge in a diagram.

Analysis (Level 3 processing) is defined as the reasoned augmentation of knowledge to generate information which is new (Marzano & Kendall, 2008). The five operations comprising analysis are also shown in Table 2. Knowledge utilisation (Level 4) processes are those used by the student to achieve a specific task (Marzano & Kendall, 2008). There are four operations within this level of mental processing. Level 5 processing, metacognition, is about monitoring, regulating and evaluating all other thought. It is also referred to as 'executive control' and comprises four operations (Marzano & Kendall, 2008). Their highest level of mental processing, the self system, determines how much energy and engagement is given to a task, and relates closely to some constructs of student motivation. It comprises four operations.

What contributes to demand?

Developers of applied (and also general) qualifications and curricula often need to know how to alter the demand of materials. We therefore think it is helpful to articulate the main contributors to demand in terms of our two selected taxonomies. It is worth reiterating that the selected taxonomies have two main dimensions: (i) knowledge domains; and (ii) levels of mental processing, which can be applied within each of the four knowledge domains (and their subdomains). Building on the published theory of the selected taxonomies' authors, we propose that there are four main methods of increasing demand in a syllabus (sometimes called a 'specification') or curriculum, or its assessment:

1. Cover a greater range of knowledge domains (or subdomains).
2. Cover higher order categories within those knowledge domains/subdomains (with the exception of the interpersonal domain, which is non-hierarchical).
3. Cover higher levels of mental processing.
4. Cover higher order operations within those levels of mental processing.

These four methods increase the *conceptual challenge* of the syllabus content, adding to both the depth and breadth of what is covered.

In addition, a fifth method of increasing demand is to increase the *volume* of content included in the syllabus, curriculum, or assessment. This is primarily another means of increasing breadth. We would argue that conceptual challenge and volume can be regarded as separate

Table 2: Summary of the six levels of mental processing and their key features.

Level of processing	*Operation*	*Description*
L6. Self system	Examining overall motivation	Identifying your level of motivation to learn particular knowledge or increase competence in a given area & then identifying the interrelationships between one's beliefs about efficacy & importance, & emotional responses that govern motivation.
	Examining emotional response	Analysing the extent to which you have an emotional response to particular knowledge & its influence on motivation.
	Examining efficacy	Examining whether you believe you have the ability, power or resource to be competent with given knowledge or at a particular skill.
	Examining importance	Examining whether knowledge is important or meets a need or personal goal.
L5. Metacognition	Monitoring accuracy	Determining the degree to which you understand given knowledge.
	Monitoring clarity	Determining the degree to which you are free from ambiguity about the knowledge.
	Process monitoring	Monitoring the success of a procedure whilst completing the procedure.
	Specifying goals	Forming clear goals and plans for accomplishing them.
L4. Knowledge utilisation	Investigating	Producing and testing hypotheses about historical, current or future events.
	Experimenting	Producing and testing hypotheses to understand physical/psychological phenomena.
	Problem solving	Trying to achieve a goal for which an obstacle is present.
	Decision making	Using knowledge to choose between alternatives.
L3. Analysis	Specifying	Constructing a new application of a known generalisation or principle.
	Generalising	Inferring new generalisations from known data.
	Analysing errors	Determining whether information is reasonable and analysing it for logic errors and inaccuracies.
	Classifying	Organising knowledge into meaningful superordinate and subordinate categories.
	Matching	Identifying similarities and differences between sections of knowledge.
L2. Comprehension	Symbolising	Creating a symbolic representation (usually an image) of the knowledge produced by integrating.
	Integrating	Refining knowledge to crucial characteristics organised in a frugal generalised form.
L1. Retrieval	Execute	Carrying out the steps in a procedure and producing a result.
	Recall	Recollecting and generating additional information.
	Recognition	Deciding whether received information is accurate, inaccurate or unknown.

dimensions of demand. Some students will find increased volume more difficult to master, whereas others will find aspects of increased conceptual challenge more difficult to master.

A method of applying the selected taxonomies to curricula

Next, we developed a method of using the selected taxonomies to check that syllabuses and other curricula (either already in existence, or in development) draw from appropriate knowledge domains. In England, the national regulator (Ofqual) requires formal assessment strategies for all regulated qualifications, and domain coverage is arguably a key aspect of validity. Awarding organisations therefore need to be able to demonstrate to the regulator that their syllabuses ('specifications') and the assessments within them draw from appropriate domains. Our method comprises five steps, which are given in Table 3.

Table 4 provides an example of how judgements made in Step 2 could be recorded in order to map the content-domain relationship (Step 3). The table should be comprehensive, covering the whole of the syllabus (or the relevant unit within it) that is being checked. To provide greater detail on the coverage of particular categories within domains of interest, the 'domain' columns could be subdivided into multiple 'category' columns.

It is worth noting that although this method was developed with applied qualifications in mind, it can also be used with general qualifications. For some subjects, it may be appropriate to exclude particular domains, instead focusing deeply on one or two domains. For example, a check of a general qualification in Physics might most usefully focus on the information and mental procedures domains only.

Demonstrating that applied qualifications cover multiple domains

To demonstrate that the selected taxonomies can be used to clarify the domain coverage of learning objectives within syllabuses and other curricula, we piloted our method with both vocational and general

Table 3: Method for conducting checks of knowledge domain coverage in syllabus and other curricula.

Step	*Details*
1. Identify the syllabus content to be checked	Checks can focus on a whole syllabus, or just one unit (or section) of the syllabus. Appropriate content is likely to include (but is not limited to): (i) learning outcomes; (ii) grading criteria.
2. Make a professional judgement about the domain(s) covered by each piece of content	Each learning outcome/grading criterion will cover one or more of the following knowledge domains: (i) Information; (ii) Mental procedures; (iii) Psychomotor; (iv) Interpersonal.
3. Map the relationship between the content and the domains	This is most easily done by creating a table to record all of the judgements made in Step 2. The completed table provides a 'mapping' of the relationship, indicating at a glance which domains are covered most and least.
4. Check for omissions and imbalances	It is important to review the table created in Step 3, to check that each unit/syllabus is covering the expected domains in the expected proportions. There is not usually any requirement for a syllabus to cover all domains, or to cover domains equally. Any omissions/imbalances could be addressed in future reviews of the qualification.
5. Write a statement of domain coverage	The statement summarises the judgements made, providing information and reassurance to end-users. For example: *"The breadth of knowledge coverage of this qualification has been reviewed using published educational taxonomies (Marzano and Kendall, 2007, 2008; Hutchins et al., 2013). Relevant knowledge and skills from the following domains are assessed: information; mental procedures; psychomotor procedures; and interpersonal knowledge and skills."*

Table 4: Example mapping of the content-domain relationship.

Learning outcome/ Grading criterion	*Unit*	*Information domain*	*Mental procedures domain*	*Psychomotor domain*	*Interpersonal procedures domain*
Example I	1			✔	
Example II	1			✔	
Example III	1				✔
Example IV	2				✔
Example V	2	✔			
Example VI	2		✔		

qualifications in a range of applied subjects, enabling us to compare their content. We obtained syllabuses ('specifications') for Cambridge Nationals (Level 2 Technical Awards targeted at 14 to 16-year-olds) in Sport Science, Sport Studies, Enterprise and Marketing, and Creative iMedia, and GCSEs in Physical Education, Business, and Media Studies. To conduct the pilot, we created a mapping table for each unit in each qualification with columns for all the 38 subcategories of knowledge and levels of mental processing. For the examination units, each exam item was typed into a separate row of the table. For the non-examination units, each sentence of the task information was typed into a separate row of the table.

Binary judgements were made as to whether each item or task sentence related to each of the categories or not. The judgements were first made by one researcher and then checked by another researcher. Judgements were based on the descriptions and explanations of the taxonomy categories in Tables 1 and 2. In order to facilitate reviewing the judgements, any aspect of the item/sentence that was judged to be related to the categories (i.e., words or phrases) was recorded in the cell of the table.

Despite the detailed guidance, the judgemental process was found to vary in difficulty across the examination items and task sentences. This was not unexpected; many studies have shown only moderate reliability of taxonomy mappings amongst both subject and non-subject experts (Coleman, 2017). Therefore, to enhance judgement consistency, regular meetings between the researchers were held. These were found to be helpful to discuss any difficult or ambiguous mappings. Also, notes were made as to how the categories had been interpreted with specific examples of words and phrases that had directed certain judgements.

Overall, this approach was deemed successful, producing mapping outcomes that cohered with experienced colleagues' perceptions of the qualifications. It was possible to conclude that the analysis revealed a different pattern of cognitive domain coverage across the two qualification types. The two Cambridge Nationals overlapped to differing degrees with GCSE content. Where there was overlap, however, the content was often assessed differently; all of the Cambridge National qualification units (bar one in each) used non-exam assessment (NEA) but their content typically overlapped with the GCSE exam component rather than its NEA component. The comparatively greater use of NEA in the Cambridge Nationals was associated with different coverage of knowledge domains and levels of mental processing compared to the GCSE. The Cambridge National NEAs focused on particular knowledge domains more than the GCSE exams did, especially mental and psychomotor procedures, and covered a wider range of levels of mental processing. For further details, see Child and Vitello (2018) and Vitello and Child (2018).

Creating a tool for writing educational objectives

Drawing from Marzano and Kendall (2007, 2008), we also explored how the selected taxonomies can best be used to write a range of types of educational objectives in new qualifications and those due to be re-developed. Shaping content in this way is preferable to checking domain coverage post-hoc, as it is better to get a qualification right first time than for it to require revisions and amendments. There are several different types of educational objective. For example, educational objectives can be: curriculum aims; syllabus aims; assessment objectives; learning outcomes; grading criteria; and detailed criteria in mark schemes.

Marzano and Kendall (2007) advocate a standard format for writing educational objectives. The authors explain that an objective has three parts:

(i) A stem;

(ii) A verb phrase (that is, the mental operation to be employed by the student);

(iii) An object of the verb phrase (that is, the knowledge that is the focus of the objective).

For example: *the student will be able to* (i.e., the stem) *present* (i.e., verb phrase) *a final proposal to a client for feedback and approval* (i.e., object of the verb phrase). Another example would be: *the student will be able to illustrate the proper hand and arm motion for the butterfly stroke*.

We found that this standard format can also be utilised in the context of Hutchins et al.'s (2013) interpersonal domain. This enabled us to create a tool for writing objectives. The tool is essentially a large table. It comprises: (i) the six levels of mental processing; (ii) the operations within them; (iii) a general form of the verb phrase for each operation; and (iv) examples of alternative verb phases which can be used to write educational objectives. The table also indicates the knowledge domains in which the operations and example verb phrases are relevant. Most operations can be used with all four domains. Additionally, examples of appropriate item types for use in assessments are also given. (An excerpt of the tool is given in Appendix A. The full tool is available from the authors.)

A key benefit of the tool is that it shows the levels of processing, the operations, and the verb phrases in a hierarchical arrangement. Figure 1 illustrates the important principle that, when writing grading criteria and learning outcomes that cover multiple levels of mental processing, this hierarchy must be adhered to rigorously. That is, higher grades must be associated with higher levels (or the same levels) of mental processing.

Where all or some grades are associated with the same level of mental processing (Examples ii and iii in Figure 1), there should be no crossover in the lines that link those grades to operations within that level of mental processing. That is, in Examples ii and iii, the hierarchy of operations within knowledge utilisation should be adhered to when selecting verb phrases for the grades.

We would argue that, prior to writing individual educational objectives, it is important to establish the desired balance of domain coverage for the whole unit or other large part of the syllabus within which the objectives lie. As mentioned previously (see Table 3), there is not usually any requirement for a syllabus to cover all domains, or to cover particular domains equally. Subjects will vary in terms of the relevance of the four domains to their content. To establish the desired balance for a new or revised syllabus, it is worth conducting market research and considering any contributory factors such as: stakeholder views (e.g., those of teachers/tutors, students, employers and higher education admissions staff); any regulatory requirements and preferences; economic trends; employment data; and in the case of vocational qualifications, national occupational standards.

We would also argue that, whilst writing educational objectives, it is important to keep a record of the domain(s) and also the level(s) of mental processing covered. This will facilitate subsequent checks that the desired balance of coverage has been achieved. It will also be useful whenever a justification of the approach taken is needed. For example, theory-driven work of this kind could make a significant contribution to the validity arguments within the assessment strategies that a qualification regulator (e.g., Ofqual) may require.

Further applications: checking item balance in an examination

Moving beyond the original aims of this study, and beyond Marzano and Kendall's (2007, 2008) suggestions, we identified two further applications of the selected taxonomies. The first of these is in the process of ensuring that an examination paper (or any other assessment) comprises the intended balance of items or marks, in terms of the domains and levels of mental processing (demand) covered.

Test design incorporates knowing what we wish to assess. Prior to creating an examination paper, it should therefore be possible to record the intended balance of items, in a simple spreadsheet for example. (This is sometimes a part of the process of creating a 'test specification' or blueprint; see Owen 2018.) To achieve this intended balance, the developer then needs to keep a record of the actual balance of items/marks requiring the use of each domain and level of processing. Once an initial draft of the examination paper is complete, this record can be compared with the record of intended balance. It is likely that the

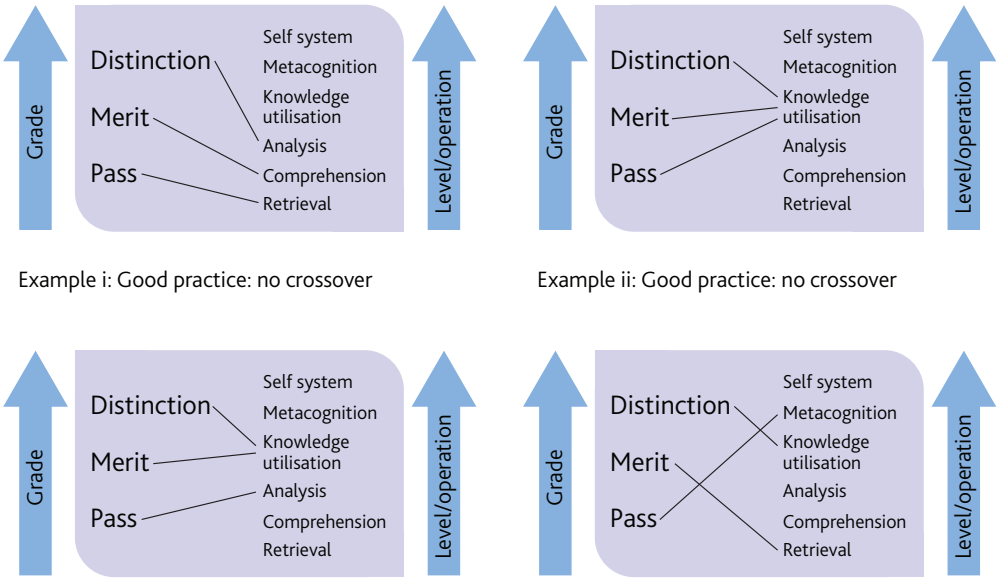


Figure 1: Examples of good and poor practice in the writing of grading criteria and learning outcomes.

percentage of items/marks in each cell of the spreadsheet of actual balance will need increasing or decreasing. To achieve this, items can be made more or less challenging by changing the command word to reflect a higher or lower level of mental processing, as needed. The wording of items can also be adjusted to alter the domain coverage. Our tool (Appendix A) will help with these processes.

Further applications: comparing and aligning content balance

Extending this suggestion, the selected taxonomies can be used to compare content balance across different types of documentation associated with education curricula, and to align them if desired. Figure 2 provides a simple illustration of this idea. Direct links between all possible pairs of document type could potentially be added to this diagram. Examinations from different years, from different awarding bodies (which may represent different countries or have a global reach), or based on different curricula, could be compared in terms of the domain coverage or demand of their content. Examination content could also be compared (and aligned) with syllabus and curriculum objectives, textbook content, and other teaching and learning resources, and these latter resources could be compared (and aligned) with one another.

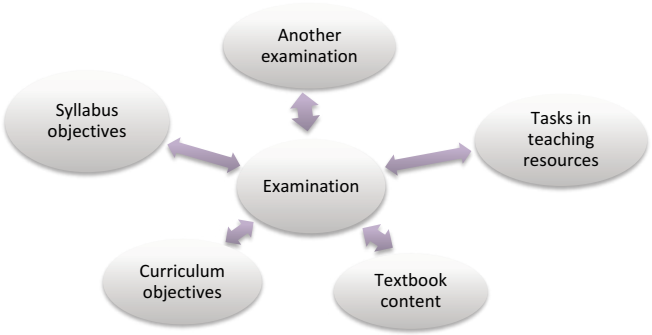


Figure 2: Potential uses of the taxonomies in comparability and alignment projects.

Conclusion

The broad aim of the present study was to explore whether any educational taxonomies designed for general education contexts could be utilised in applied educational contexts, and we have demonstrated this to be the case. As with all studies, there were several limitations to the work undertaken. Of the hundreds of educational taxonomies in the public domain, only nine could be reviewed systematically in the time available. Moreover, the piloting of our mapping method indicated that its judgemental process varied in difficulty across examination items and task sentences. Whilst not unexpected, this necessitated regular discussions between those applying the method. Nonetheless, the two selected taxonomies were used successfully with qualifications in multiple applied subjects.

A key question arising from this study is that of how the selection and application of educational taxonomies relates to validity, which is a hallmark of quality for educational measurement. In the academic literature, there are many conceptions of validity (or of multiple types or subtypes of validity). These conceptions are evolving constantly since they are contested continuously by theorists. For a detailed discussion, see Newton and Shaw (2014). Rather than embroiling qualifications developers in this complex debate by exploring validity theoretically, we propose that within the context of this study and its applications, it is more beneficial to take a pragmatic approach. In common parlance, the 'validity' of an assessment is often taken to mean its 'authenticity' or 'integrity'. That is, does it assess what it purports to assess? Demonstrating that the content of a course and its associated assessments cover what stakeholders require is a means of demonstrating validity in this sense. Regulated qualifications in England usually require an assessment strategy which includes a validity argument; a mapping of domain coverage and levels of mental processing can make a valuable contribution to this.

Arguably, taxonomies of educational objectives are underused at present. They have the potential to add rigour to multiple aspects of qualifications and curriculum design and development. We have shown in this study that our selected taxonomies (Marzano and Kendall, 2007, 2008; Hutchins et al., 2013) can provide developers with a concrete means to demonstrate to stakeholders the domain coverage of applied qualifications and curricula. In addition to supporting such development, the selected taxonomies could improve ongoing processes for monitoring, comparing and aligning the functioning of existing qualifications, both applied and general, within and across awarding organisations.

References

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*. Longman.

Atkinson, S. P. (2013). Taxonomy Circles: Visualizing the possibilities of intended learning outcomes. *Learning and Teaching Working Paper*. https://spatkinson.files.wordpress.com/2015/01/taxonomy-circles-atkinson-aug13.pdf

Bandaranaike, S., & Willison, J. (2015). Building capacity for work-readiness: Bridging the cognitive and affective domains. *Asia-Pacific Journal of Cooperative Education, Special Issue, 16*(3), 223–233.

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning—the SOLO taxonomy*. (1st ed.). Academic Press.

Bloom, B. S., Engelhart, M. D., Furst, E. D., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: the classification of educational goals. Handbook 1 Cognitive Domain*. Longmans.

Carpenter, T. D., & Wisecarver, M. M. (2004). *Identifying and Validating a Model of Interpersonal Performance Dimensions*. Alexandria: Army Research Institute for the Behavioral and Social Sciences.

Child, S. F. J., & Vitello, S. (2018, September). *An exploration of domain coverage of Cambridge National and GCSE qualifications*. Paper presented at the annual conference of the British Education Research Association, Northumbria University, Newcastle, UK.

Coleman, V. (2017). On the reliability of applying educational taxonomies. *Research Matters: A Cambridge Assessment publication, 24*, 30–37.

Dave, R. (1967). *Psychomotor domain*. Paper presented at the International Conference of Educational Testing, Berlin.

Gleeson, J., & O'Flaherty, J. (2013). School-based initial vocational education in the Republic of Ireland: the parity of esteem and fitness for purpose of the Leaving Certificate Applied. *Journal of Vocational Education and Training, 65*(4), 461–473.

Harrow, A. J. (1972). *A Taxonomy for the Psychomotor Domain: A Guide for Developing Behavioural Objectives*. David McKay.

Hauenstein, A. D. (1998). *A conceptual framework for educational objectives: A holistic approach to traditional taxonomies*. University Press of America.

Hutchins, S., McDermott, P., Carolan, T., Gronowski, M., Fisher, A., & DeMay, M. (2013). Interpersonal Skills Summary Report *Research Note 2013–03*. Alion Science and Technology, Inc.

Kämäräinen, P., & Fischer, M. (2008). Research on Technical and Vocational Education and Training (TVET) in the Context of European Cooperation. In F. Rauner & R. MacLean (Eds.), *Handbook of Technical and Vocational Education and Training* (pp. 135–143). Springer Science+Business Media B.V.

Kane, M. (1998). Choosing Between Examinee-Centered and Test-Centered Standard-Setting Methods. *Educational Assessment, 5* (3), 129–145.

Klein, C., DeRouin, R. E., & Salas, E. (2006). Uncovering Workplace Interpersonal Skills: A Review, Framework, and Research Agenda. In G. P. Hodgkinson & J. K. Ford (Eds.), *International Review of Industrial and Organizational Psychology 2006* (Vol. 21). John Wiley & Sons.

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory into Practice, 41*(4), 212–218.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of Educational Objectives, Handbook II: Affective domain*. David McKay.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1999). *Taxonomy of Educational Objectives Handbook II/ Affective Domain* (2nd ed.): Longman Publication Group.

Laxmisan, A., Hakimzadaa, F., Sayan, O. R., Green, R. A., Zhang, J., & Patel, V. L. (2007). The multitasking clinician: Decision-making and cognitive demand during and after team handoffs in emergency care. *International Journal of Medical Informatics, 76*, 801–811.

Marzano, R. J., & Kendall, J. S. (2007). *The New Taxonomy of Educational Objectives* (2nd ed.). Corwin Press Inc.

Marzano, R. J., & Kendall, J. S. (2008). *Designing & assessing educational objectives; applying the new taxonomy*: Corwin Press Inc.

McGrath, S., Akoojee, S., Gewer, A., Mabizela, M., Mbele, N., & Roberts, J. (2006). An examination of the vocational education and training reform debate in Southern Africa. Compare: *A Journal of Comparative and International Education, 36*(1), 85–103.

Newton, P. E., & Shaw, S. D. (2014). *Validity in Educational and Psychological Assessment*. Sage.

Owen, N. (2018). Test Specifications. In J. I. Liontas, T. International Association and M. Delli Carpini (Eds.), *The TESOL Encyclopedia of English Language Teaching*. Wiley Blackwell.

Savic, M., & Kashef, M. (2013). Learning outcomes in affective domain within contemporary architectural curricula. *International Journal of Technology and Design Education, 23*(4), 987–1004.

Vitello, S., & Child, S. (2018, July). *Vocational qualifications for 14–16 year olds: Exploration of knowledge, skills and teacher perceptions*. Paper presented at the 5th International Conference on Employer Engagement in Education and Training, London, UK.

Appendix A: Excerpt from tool for writing objectives, based on Marzano and Kendall's taxonomy of educational objectives

All learning outcomes and grade criteria follow the standard form: "The learner is able to", then a verb phrase, then an object of the verb phrase.

Level	*Operation*	*Verb phrase*		*Relevant knowledge domains*				*Item type*	*Source*
		General form of the verb phrase	*Marzano and Kendall's examples of alternative verb phrases*	*Information*	*Mental procedures*	*Psychomotor*	*Interpersonal knowledge/ skills*		
L1. Retrieval	Recognising	Validate the accuracy of … they are given	Recognise Select from a list Identify from a list Determine if the following statements are true	✔	✔	✔	✔	Forced choice	(Marzano and Kendall 2008, p.42)
	Recalling	Generate accurate information…	Recall Exemplify Name List Label State Describe	✔	✔	✔	✔	Short constructed response Cloze	(Marzano and Kendall 2008, p.42)
	Executing	Carry out a…	Add Subtract Multiply Divide Apply Demonstrate Draft Complete Make Solve Read Use Write		✔	✔	✔	Short constructed response with execution of the procedure Forced choice (for Mental tactics algorithms and simple rules)	(Marzano and Kendall 2008, p.42

Cambridge Assessment Network

Take your assessment ambitions further with our range of online CPD accredited courses. With multiple start dates throughout the year, our 9-week online courses can be taken at a time to suit you.

A101: Introducing the Principles of Assessment

Understand what's behind good assessment with a detailed grounding in the principles of assessment.

A102: Introducing Assessment Practice

Learn how exams are created and marked as you discover the key stages of the assessment cycle.

A103: Introducing Data Literacy

Make greater use of assessment data and understand the basics of managing, reading and interpreting data.

A104: Psychometrics in Educational Assessment

Expand what you know about assessment analysis in order to ensure tests are reliable, valid and fair.

Find out more at: cambridgeassessment.org.uk/online

Cambridge Assessment Network

Training for your workplace

Flexible assessment learning packages

Whatever your organisation's professional development needs, as assessment leaders we are committed to supporting your people.

Our ready-made options cover most requirements giving you an opportunity to put learning into your own context. Example sessions include:

- The purposes and principles of good assessment
- Writing multiple choice questions for impact
- Writing effective examination assessment questions

We also develop bespoke training solutions for awarding bodies, educational organisations, professional associations, universities, the private sector and government departments.

Just let us know what you need.

> "The session gave participants ideas, strategies and some reflective learning as take aways as well as some useful aide memoire handouts. The Network team did an excellent job and covered our brief to the letter."
>
> Dr Andrew Harries, University of Central Lancashire, on an in house Multiple Choice Questions session

Find out more at:
cambridgeassessment.org.uk/your-cpd

Research News

Anouk Peigne Research Division

As the coronavirus pandemic took hold, most education and assessment conferences around the world were cancelled. Throughout the spring and summer of 2020 a significant number of our researchers became involved in developing and testing the extraordinary changes to systems and processes that were necessary to award grades to candidates in an exam series in England where no students were taking actual exams. Events moved extremely quickly, so we felt it was important to maintain an overview and a record of what was happening at national level in the different parts of the UK. Our 'Covid-19 Curriculum Watch' extended blogs (see 'Insights' below) kept track of some of the main events in policy, teaching and assessment during the period.

Conference Presentations

Society for Research into Higher Education (SRHE)

The annual SRHE International Conference on Research into Higher Education took place in December 2019 in Newport, Wales, and the themes were *Creativity, Criticality and Conformity in Higher Education*.

Carmen Vidal Rodeiro presented *How does A-level subject choice and students' background characteristics relate to Higher Education participation?*

The International Academic Forum (IAFOR) 8th European Conference on Education

This conference took place remotely from London in July 2020. The sessions were aligned with the challenges faced as a result of Covid-19.

Emma Walland presented *Remote marking of high-stakes examinations: leadership, challenges and strategies.*

Insights

Insights provides a platform for sharing Cambridge Assessment's views and research on the big education topics that impact assessment around the globe.

The following Insights have been published by the Research team since *Research Matters*, Issue 29:

Cambridge Assessment Education and Curriculum Research Team. (2020, 2 July). Covid-19 Curriculum Watch: Education Policy in the first 3 months of the pandemic. Cambridge Assessment Insight. https://www.cambridgeassessment.org.uk/insights/uk-education-policy-during-covid-19-pandemic-topic-one/

Cambridge Assessment Education and Curriculum Research Team. (2020, 14 July). Covid-19 Curriculum Watch 2: Changes to assessment in response to the pandemic. Cambridge Assessment Insight. https://www.cambridgeassessment.org.uk/insights/uk-changes-to-assessment-during-covid-19-pandemic-topic-two/

Cambridge Assessment Education and Curriculum Research Team. (2020, 6 August). Covid-19 Curriculum Watch 3: Learning access, resources, and assessment guidance during the pandemic. Cambridge Assessment Insight. https://www.cambridgeassessment.org.uk/insights/learning-access-resources-and-assessment-guidance-during-covid-19-pandemic-topic-three/

Cambridge Assessment Education and Curriculum Research Team. (2020, 26 August). Covid-19 Curriculum Watch 4: Curriculum choices and the pandemic. Cambridge Assessment Insight. https://www.cambridgeassessment.org.uk/insights/curriculum-choices-and-the-pandemic-curriculum-watch-topic-four/

Publications

The following reports and articles have been published since *Research Matters*, Issue 29:

Bramley, T. (2020). Metaphors and the psychometric paradigm. *Assessment in Education: Principles, Policy & Practice*, 27(2), 178–191. https://www.tandfonline.com/doi/full/10.1080/0969594X.2020.1731421

Constantinou, F. (2020). Examination questions as a form of communication between the examiner and the examinee: a sociolinguistic perspective on assessment practice. *Cambridge Journal of Education* (ahead of print). https://www.tandfonline.com/doi/abs/10.1080/0305764X.2020.1761293

Gill, T. (2019). *The relationship between local authority schools' funding levels and their provision of qualifications*. Cambridge Assessment Research Report. Cambridge, UK: Cambridge Assessment. https://www.cambridgeassessment.org.uk/Images/the-relationship-between-local-authority-schools-funding-levels-and-their-provision-of-qualifications.pdf

Johnson, M., & Coleman, V. (2020). Out of their heads: using concept maps to elicit teacher-examiners' assessment knowledge. *International Journal of Research & Method in Education* (ahead of print). https://www.tandfonline.com/doi/full/10.1080/1743727X.2020.1804542

Macinska, S., & Benton, T. (2020). *The usefulness of detailed marks within the levels of levels-based mark schemes*. Cambridge Assessment Research Report. Cambridge, UK: Cambridge Assessment. https://www.cambridgeassessment.org.uk/Images/593879-the-usefulness-of-detailed-marks-within-the-levels-of-levels-based-mark-schemes.pdf

Vidal Rodeiro, C. L., & Vitello, S. (2019). *Vocational Qualifications at Key Stage 4 and Key Stage 5: who takes them and how they fit into students' programmes of study*. Cambridge Assessment Research Report. Cambridge, UK: Cambridge Assessment. https://www.cambridgeassessment.org.uk/Images/579348-vocational-qualifications-at-key-stage-4-and-key-stage-5-who-takes-them-and-how-they-fit-into-students-programmes-of-study.pdf

Williamson, J., & Child, S. (2020). Mark scheme design for school- and college-based assessment in VTQs. *Journal of Vocational Education & Training* (ahead of print). https://www.tandfonline.com/doi/full/10.1080/13636820.2020.1782454

UNICEF Learning Passport programme

The Learning Passport programme includes a pan-Cambridge University collaboration with UNICEF, as well as working with Microsoft. An aim of the Learning Passport programme is to offer displaced children the opportunity to continue education anywhere in the world, using quality resources.

The following reports have been published by Cambridge:

Cambridge Assessment & the University of Cambridge (2020). *The Learning Passport: Curriculum Framework Adaption Guidance Making progress possible: Improving the quality of education for vulnerable children everywhere*. Cambridge, UK: Cambridge University Press. https://www.cambridge.org/files/4415/8465/3367/The_Adaptation_Guidance.pdf

Cambridge University Press & Cambridge Assessment (2019). *The Learning Passport Research and Recommendations Report: Summary of Findings*. Cambridge, UK: Cambridge University Press & Cambridge Assessment. https://www.cambridgeassessment.org.uk/Images/577273the-learning-passport-research-and-recommendations-report..pdf

Oates, T., Johnson, M., Fitzsimons, S., Coleman, V., & Greatorex, J. (2020). *The Learning Passport: Curriculum Framework (Maths, Science, Literacy)*. Cambridge, UK: Cambridge Assessment. https://www.cambridge.org/files/7615/8465/3386/The_Curriculum_Framework-Maths_Science_Literacy.pdf

The University of Cambridge, Cambridge University Press & Cambridge Assessment (2020). *The Learning Passport: Research and Recommendations Report*. Cambridge, UK: Cambridge University Press. https://www.cambridge.org/files/8615/8465/3596/The_Research_and_Recommendations_Report.pdf

Sharing our research

We aim to make our research as widely available as possible. Listed below are links to the places where you can find our research online.

- *Journal papers and book chapters:* www.cambridgeassessment.org.uk/our-research/all-published-resources/journal-papers-and-book-chapters/
- *Research Matters* (in full and as PDFs of individual articles): www.cambridgeassessment.org.uk/research-matters
- *Conference papers:* www.cambridgeassessment.org.uk/our-research/all-published-resources/conference-papers/
- *Research Reports:* www.cambridgeassessment.org.uk/our-research/all-published-resources/research-reports/
- *Data Bytes:* https://www.cambridgeassessment.org.uk/our-research/data-bytes/
- *Statistics reports:* https://www.cambridgeassessment.org.uk/our-research/all-published-resources/statistical-reports/
- *Blogs*: www.cambridgeassessment.org.uk/blogs/
- *Insights* (a platform for sharing our views and research on the big education topics that impact assessment around the globe): www.cambridgeassessment.org.uk/insights/
- Our *Youtube* channel: www.youtube.com/user/CambridgeAssessment1 contains Research Bytes (short presentations and commentary based on recent conference presentations), our online live debates #CamEdLive, and Podcasts.

You can also learn more about our recent activities from Facebook, Instagram, LinkedIn and Twitter.

Contents / *Issue 30* / Autumn 2020

Cambridge Assessment
The Triangle Building
Shaftesbury Road
Cambridge
CB2 8EA
United Kingdom

+44(0)1223 553985
ResearchDivision@cambridgeassessment.org.uk
www.cambridgeassessment.org.uk

7867929311

ISSN: 1755 6031