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

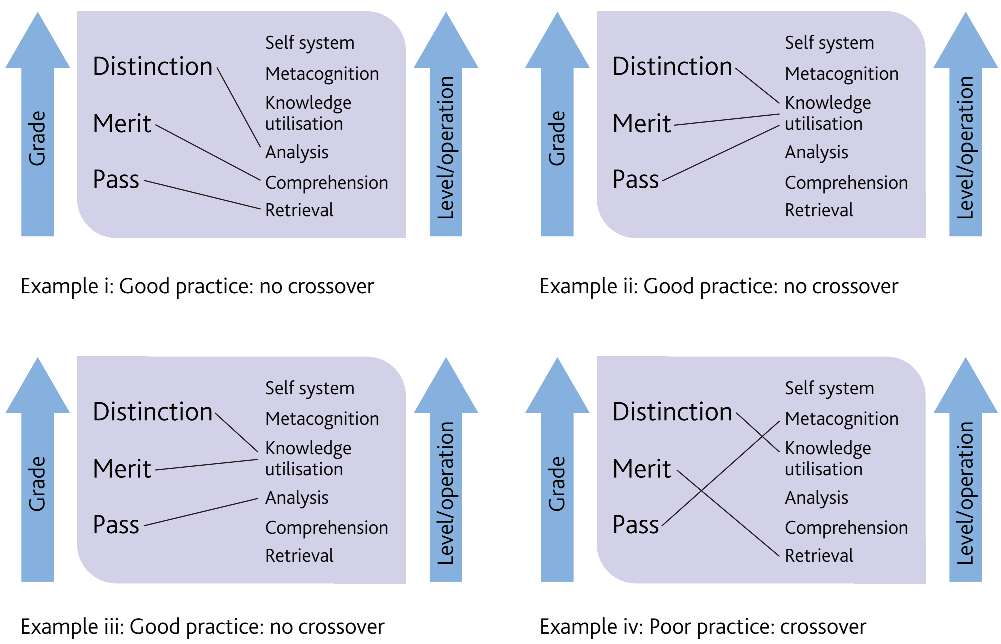

Figure 1: Examples of good and poor practice in the writing of grading criteria and learning outcomes.

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.

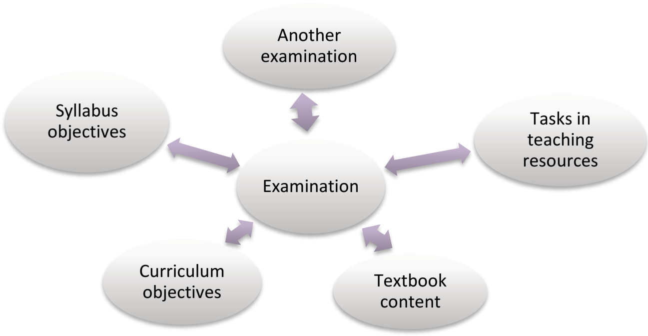

Figure 2: Potential uses of the taxonomies in comparability and alignment projects.

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… 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