

This paper was downloaded from

The Online Educational Research Journal
(OERJ)

www.oerj.org

OERJ is an entirely internet-based educational research journal. It is available to anyone who can access the web and all articles can be read and downloaded online. Anybody can submit articles as well as comment on and rate articles. Submissions are published immediately provided certain rules are followed.

Do experts' views of specification demands correspond with established educational taxonomies?

Abstract

The aim of this study was to develop a framework for comparing the specification demands of cognate units from different types of qualifications. Experts from three subjects provided views about the demands of cognate units from various types of qualifications. Educational taxonomies in the Affective/Cognitive/Interpersonal/Metacognitive/Psychomotor domains were used in coding experts' views of specification demands. The coding indicated whether their views corresponded to part of the taxonomies. A reasonable proportion of the data corresponded with at least one taxonomy. The taxonomies were relevant to the subjects and qualification types used in the research. Therefore, the taxonomies could be used as a framework for comparing the demands of cognate units from different types of qualifications. The taxonomies are based on more substantial research than the awarding body rating scales in comparability studies; therefore they could provide a more robust framework for specification demands comparisons in comparability studies.

Introduction

Comparability is an important educational research topic in the UK¹. Qualifications are compared to ensure the comparability of demand of pathways to employment, study programmes or training. Therefore, the aim of this study was to develop a re-usable framework for comparing the specification demands of cognate units from different types of qualifications used in comparability research in the UK. In this article, the UK qualifications system is described, terms defined, and comparability studies are detailed before focusing on the present study.

Awarding bodies in the UK offer qualifications which are chosen by centres (for example, schools) and learners. Educational providers and learners have a choice between different types of qualifications and awarding bodies. They may include:

- General qualification (GQ)
- Vocational qualification (VQ)
- Vocationally-related qualification (VRQ)
- Principal learning (PL)

GQs are usually academic qualifications studied in schools and further education colleges. For instance, General Certificates of Secondary Education (GCSE) are GQs assessed using paper-based examinations near the end of compulsory schooling in England.

VQs (sometimes referred to as occupational or technical qualifications) are usually related to a job or occupation. They are often taken within a company. Examples include National Vocational Qualifications (NVQs) for jobs such as Hairdressing, Printing or Plumbing.

NVQs are derived from the national occupational standards (NOS) that state what a competent person in a job is expected to do. NVQ learners may already be in a job and seeking formal accreditation for their job competence or they could be studying and working part-time. Learners' competence is generally assessed on the job.

VRQs are intended to provide a broad education for further training, education or entering employment within an occupational sector. VRQs are generally studied in schools or further education. The mode of

¹ Note that in this report, the UK is used to refer to England, Wales and Northern Ireland. Scotland is not included as it has separate qualifications.

assessment varies greatly and includes practical assessments, paper/computer-based tests and examinations. They are available in areas such as Health and Social Care, IT Users and Customer Service.

VQ, VRQ and GQ are broad categories. PL qualifications are intended to equip learners with a balanced mix of knowledge, understanding and skills in their chosen field of study (Office of Qualifications and Examinations Regulation, 2011). They are designed to prepare learners for employment in an occupational sector or further study in education. They are available in subjects such as: Engineering; Construction and the Built Environment; and Hair and Beauty. PL was originally a compulsory part of diplomas, which were composite qualifications where learners also took units from other qualifications such as GCSEs and NVQs. Diplomas were often mistaken for vocational qualifications (Ertl and Stasz, 2010). Diplomas were discontinued in 2013, but PL is still available as a qualification. For a detailed profile of state funded qualifications and assessment in England see Isaacs (2010).

Terms and definitions

It is important to clarify some terms and concepts before discussing comparability research in more detail.

A *taxonomy* is

...a classification system that establishes the hierarchy of the parts to the whole. An orderly classification of items according to their presumed relationship. A hierarchy presumes that a lower order class is a prerequisite to a higher order class. Also in a hierarchy, a higher order class incorporates more specific lower order classes (Hauenstein, 1998, 2).

Demands are the required level of knowledge, skill and understanding. The concept of demands is applied to many aspects of general and vocational education and training as well as in the world of work. The literature refers to the demands of:

- An examination question or item (Pollitt, Ahmed, and Crisp, 2007)
- An examination question paper (Pollitt et al., 2007)
- A task such as an activity that forms part of a job (Garrett, 2010; Laxmisan, Hakimzadad, Sayan, Green, Zhang, and Patel, 2007) or a classroom activity or similar [REDACTED]
- A specification² [REDACTED]
- A curriculum or part of a curriculum (Adey and Shayer, 1994; Shayer, 1978)

The focus of the present research was *specification demands*, which are a complex area needing some illumination. Pollitt et al. (2007) explained that demands are intended inputs, that is, demands are what the specification writers/revisers intend to require of typical learners. They noted that demands are qualitative, and therefore their size or educational significance cannot be reliably measured. Furthermore, even quantitative analyses of demands are relative comparisons rather than absolute measures.

An *expert* is a high ranking assessor, verifier or moderator contracted by an awarding body to assure the quality of performance is appropriate, often by checking the assessment judgements made by others. Alternatively, an expert can be someone who revises specifications and is contracted by the awarding body. Revisers, senior revisers, assessors, verifiers and moderators are experts in their subject or vocational specialism.

Performance evidence might be examination scripts, an artefact the learner crafted, a video of a drama in which the learner acted, a witness testimony from a line manager describing how a learner met an assessment criteria, a learner leading a meeting, and so on.

² A *specification* is "The complete description - including optional and mandatory aspects - of the content, assessment arrangements and performance requirements for a qualification. A subject specification forms the basis of a course leading to an award or certificate. Formerly known as a 'syllabus'". (Qualifications and Curriculum Development Authority, 2010)

Comparability research methods

Comparability is a major issue in educational research in the UK. Awarding bodies and the Office of Qualifications and Examinations Regulation³ compare standards of qualifications between awarding bodies, over time and between cognate qualifications at the same level. A key rationale for undertaking such work is to avert inequalities which would be formed by the existence of more/less demanding (challenging) pathways to employment, study programmes or training.

Several different methods have been used in past comparability research. Ideally comparisons are made in terms of:

- Task demands
- Specification demands
- The quality of learners' performance.

Examples of comparability studies using one or more of the above are available on a compact disc as an appendix to Newton, Baird, Goldstein, Patrick, and Tymms (2007). In some studies, such as Fearnley (2000) and Guthrie (2003), the comparison of task demands and specification demands was interwoven rather than separated into two distinct parts of the research. However, other comparability studies, for example Office of Qualifications and Examinations Regulation (2012), did separate task demands from specification demands.

To conduct robust comparability research about task demands either all the tasks or a sizeable and representative sample of the tasks are required. Likewise, for sound comparability research about quality of performance, a sizeable and representative sample of performance evidence is required. In the case of VQs, VRQs and new qualifications, there can be a lack of representative samples of tasks, performance evidence, centres or learners. In such circumstances comparability research is restricted to specification demands.

Previous studies which investigated specification and task demands, for example, Gray (2000) and Pritchard, Jani, and Monani (2000), tended to use Kelly's repertory grid technique (KRG) in their research design. KRG is a specialist type of interview used in a variety of contexts for systematically recording how people view the world. The interview is designed to elicit the constructs which are the means by which people evaluate elements (objects, events, people, ideas and so on) that make up their world. These constructs are bipolar, that is the constructs can be summarised as a phrase or word and its converse, for example, *positive - negative, high achieving – low achieving, holds my attention – leaves me uninterested*. To elicit these constructs, research participants' answers to questions such as the following were recorded.

- Which two elements are similar and different from a third?
- How are these two elements similar and how are they different from the third?
- Can you summarise the similarity into a short phrase?
- Can you summarise the difference into a short phrase?

For more details of KRG see Bannister (1965), Cohen and Manion (1994) or Fransella, Bell, and Bannister (2004).

KRG formed part of the overall design in many comparability studies (such as Gray (2000), Pritchard et al. (2000), Arlett (2003), Edwards and Adams (2002)) as follows:

1. Kelly's repertory grid interviewing was used to consult one expert from each qualification about demands. Ideally, the experts had a senior level of responsibility for one of the

³Office of Qualifications and Examinations Regulation (Ofqual) is the current national regulator of awarding bodies in England. QCA (Qualifications and Curriculum Authority), QCDA (Qualifications and Curriculum Development Agency), NCVQ (National Council for Vocational Qualifications) and SCAA (School Curriculum and Assessment Authority) were the predecessor organisations.

- qualifications in the comparability study. The views from the experts were collated to form a master list of demands with no duplicates.
2. A questionnaire was written constituting a series of demands each based on the interview data. Rating scales were added to the demands in the series. One end of each scale was the most demanding and the other end was the least demanding.
 3. A wider group of experts was asked to rate each specification and associated assessment materials on each rating scale.
 4. The ratings were analysed to compare the demands of the qualifications.

There are three important challenges of this approach:

1. Experts sometimes talked about characteristics of specifications which are not demands. Pollitt et al. (2007) suggested that these characteristics should be excluded from the list of demands, so that they are not rated by the wider group of experts.
2. The number of experts who were interviewed was often small (see Table 1). Therefore the list of demands used in the comparability research was based on limited data.
3. The list of demands on which experts rate specifications was often specific to a subject and a small number of qualifications and could not be easily re-used. Therefore a new scale was developed for each research study. Pollitt et al. (2007) called for the development of a demands framework that could be used in all comparability studies removing the need for the elicitation phase.

Table 1: Awarding body studies of task and specification demands

Awarding body research study	Number of experts who provided demands
Adams and Pinot de Moira (2000)	6
Fearnley (2000)	4
Gray (2000)	6
Pritchard <i>et al.</i> (2000)	6
Arlett (2002)	3
Arlett (2003)	3
Guthrie (2003)	8
████████████████████	5
████████████████████	6
Edwards and Adams (2002, 2003)	6

An alternative approach is to use taxonomies for the comparisons. There are several established taxonomies which were designed for classifying educational objectives, as well as aiding curriculum and test development. Of 194 comparability studies carried out by Office of Qualifications and Examinations Regulation or awarding bodies, only 19 referred to established taxonomies. Examples include Johnson and Cohen (1983) and Willmott (1980). Several studies (e.g. Qualifications and Curriculum Authority (2007a, 2007b, 2008)) used CRAS, which is primarily a cognitive taxonomy for comparing the demands of individual examination questions (Pollitt et al., 2007; Pollitt, Hughes, Ahmed, Fisher-Hoch, and Bramley, 1998).

It has been claimed that using taxonomies in comparability studies may improve comparability methods (Coles and Matthews, 1995; Coles and Matthews, 1998; Pollitt et al., 2007). Taxonomy use might overcome challenges with the KRG method listed above. In contrast to the current awarding body approach, some existing taxonomies provide an indication of what is more or less demanding without referring to additional characteristics of specifications, such as the clarity of the mark scheme. They are often based on a considerable amount of research. In addition, they are largely independent of subjects and type of education (academic or vocational).

However, there are potential pitfalls with using established taxonomies in comparability studies. It could be claimed that some taxonomies (e.g. Anderson, Krathwohl, Airasian, Cruikshank, Mayer, Pintrich, Raths, and Wittrock (2001), Bloom, Engelhart, Furst, Hill, and Krathwohl (1956), Harrow (1972), Hauenstein (1998), Krathwohl, Bloom, and Masia (1964), and Simpson (1966)) are not necessarily appropriate to current UK qualifications as they were developed in a different context (often the curriculum and testing situation in the US). Other taxonomies, such as CRAS, only cover cognitive demands.

There are broadly five main domains of demands:

- *Affective*, that is the “knowledge involved with developing dispositions (prevailing tendencies) in relation to feelings, values, and beliefs.” (Hauenstein, 1998, p3)
- *Cognitive*, that is the “knowledge involved with the process of knowing and development of intellectual skills and abilities.” (Hauenstein, 1998, p3)
- *Interpersonal*, which refers to positive relationships between people.
- *Metacognitive*, that is the “awareness and conscious use of the psychological processes involved in perception, memory, thinking and learning” (Moseley *et al.*, 2004, p62)
- *Psychomotor*, that is the “knowledge involved in developing physical abilities and skills following an input of information/content.” (Hauenstein, 1998, p3).

Each of these domains is important in the area of vocational education, training, qualifications and employability (Berger, Kipfer, and Buchel, 2008; Gunta, 2013; Marope, Chakroun, and Holmes, 2015; Munby, Zanibbi, Poth, Hutchinson, Chin, and Thornton, 2007).

Research aim

As discussed, there were strengths and weaknesses in previous comparability studies using KRG or taxonomies. However, it was interesting to find out whether combining the two methods would overcome some of the limitations associated with previous comparability research about demands. The strategy for the research was to relate experts’ views of demands with established taxonomies. If there was correspondence between the taxonomies and the experts’ views of demands then the taxonomies could be used to develop a re-usable framework for comparing the specification demands of cognate units from different types of qualifications in future comparability research. This process paralleled the approach taken to develop CRAS (Hughes, Pollitt, and Ahmed, 1998).

Therefore, the research question was:

Do experts’ views of specification demands correspond with established educational taxonomies?

Method

Choosing taxonomies

Taxonomies were used in the study to give a more precise definition of demands than the broad definition provided by the interviewees and to indicate what was more and less demanding. Several different taxonomies were investigated for use in the coding, not all of which proved suitable for use in the present research.

For the five domains of demand, taxonomies were chosen that:

- Were accessibly written
- Were applicable to education for adolescents and adults
- Were applicable to academic, work based and practical learning
- Included some differentiation between what is more and less demanding.

Taxonomies that were divided into several levels of demand were identified for four of the domains. The final taxonomies used were Hauenstein’s (1998) Affective, Cognitive and Psychomotor taxonomies, and Howell and Caros’ (2006) Metacognitive taxonomy. No taxonomy was found for the interpersonal domain that indicated what was more and less demanding. Rackham and Morgan’s (1977) Interpersonal taxonomy was used despite its lack of hierarchically ordered categories.

KRG Elements

The elements were cognate units from different types of qualification offered by OCR⁴ in each of three subjects. The name of the PL qualification was used as the name for the subject. There were no cognate VQ units available for Creative and Media, or Engineering. The details of the units are given in Table 2.

Table 2: Units used in the interviews were from the following subjects, qualifications and qualification types

Subject	Number of units	Qualification type
Creative and Media	1	PL
	1	GQ
	1	VRQ
Engineering	1	PL
	2	GQ
	1	VRQ
Society, Health and Development	1	PL
	2	VRQ
	1	GQ
	2	VQ

Specifications are substantial documents with details of several units as well as administrative information such as enquiries about results and the overlap with other qualifications. Therefore only relevant extracts of specifications were used as the elements. The extracts contained the following information:

- The aims and objectives
- The unit content
- The grade or performance descriptors (if applicable)
- The assessment and qualification structure
- Information about guided learning hours and length of assessments
- Teaching arrangements.

Participants

Twelve experts were recruited, four from each subject. At least one expert was recruited from each qualification type in the subject (except Engineering VRQ). The experts represented the units/qualifications at a senior level (generally Principal Assessor or above) and several had experience of more than one unit. The sample of experts was designed to include:

- A variety of types of qualifications and subjects
- Members of the target population of experts who might be participants in future research about the comparability of demands.

Procedure

Prior to the KRG interviews the experts were briefed on the interview structure and interview topic. They were provided with a broad definition of specification demands and a definition of each domain and given time to familiarise themselves with the specification extracts.

The interview schedule was devised from advice in Jankowicz (2004). It began by re-clarifying the topic as specification demands, broadly defining specification and specification demands, and describing the

⁴ Oxford Cambridge and The Royal Society for the encouragement of Arts, Manufactures and Commerce. An awarding body.

domains. Experts were then familiarised with making comparisons about specification demands. They were asked to:

- Consider a triad of units
- Say which two units were the same in a demand domain (specified by the interviewer) and how the third was different
- Summarise the similarity and difference in two short statements (poles)
- Say which (if either) statement described something which was more demanding.

The process was repeated until the experts could not identify any further similarities and differences.

The team of four interviewers conducted practice interviews with colleagues to familiarise themselves with the interview technique and to standardise the interview procedure. Subsequently, each interview with the experts was conducted by two interviewers, where one interviewer asked questions and the other recorded the poles. This was a departure from usual KRG interviews which are conducted by one researcher. However, there are exceptions, such as Suto and Nádas (2009), where two researchers conducted each interview.

Sometimes experts' comments were not based on the triad technique and the similarities and differences were derived from between zero and six units. Nevertheless, the comments were recorded as bipolar constructs.

Nine interviews were undertaken face to face and three interviews were conducted over the telephone.

Analysis

A total of 361 bipolar constructs were elicited in the KRG interviews (see Table 3 for examples).

Table 3: Examples of bipolar constructs

Subject	Pole 1 (similarity between units)	Pole 2 (how another unit was different)
Creative and Media	Doing what you want to do	Working with others
Engineering	Require the candidates to visit the real world and interact with clients and the public	Swapping of knowledge within a confined group
Society, Health and Development	Less prescriptive assessment open to interpretation	Prescriptive assessment

Each pole of the bipolar construct was coded jointly by three researchers from the team of interviewers. First the researchers decided which taxonomies a pole corresponded to (for example, Cognitive, Affective etc.). Then they decided which level within the taxonomy the pole corresponded to (for example, Cognitive level 1, Affective level 5). Note that each pole could correspond to several taxonomies, but only one level within a particular taxonomy. Any differences in views about the appropriate code for a pole were discussed and a consensus was reached among the researchers before the final decision was recorded.

Once all the poles were assigned the researchers revisited the data to confirm the coding decisions. Table 4 contains examples of poles with their corresponding domain, taxonomy level and category. It is important to note that the coding did not use the taxonomies to estimate demands, which is often the intended use of taxonomies; rather it indicated that the expert's view corresponded with part of that taxonomy. Note also that the taxonomy descriptions used for coding were fuller than the summary descriptions given in this table.

Table 4: Examples of poles with the corresponding domain and taxonomy category

Example of a pole	Subject	Domain of demand	Taxonomy Category
<i>Being aware of the things you think are prejudiced and knowing how you become prejudiced</i>	Society, Health and Development	Affective	“1.0 Receiving. Disposition to be aware, willing and attentive" (Hauenstein, 1998, p80)
<i>Candidates are required to clearly and effectively communicate their decision making process through written communication</i>	Creative and Media	Cognitive	“2.0 Comprehension. Ability to translate and interpret ideas, and extrapolate content information". (Hauenstein, 1998, p53)
<i>Collaboration with peers and employers requires effective unfamiliar interpersonal skills</i>	Society, Health and Development	Interpersonal	Behaviours include Supporting, Disagreeing, Defending/attacking, Blocking/difficulty stating and Testing understanding (Rackham and Morgan, 1977)
<i>Involves talking to clients and people outside of school environment</i>	Creative and Media		
<i>Candidates are required to analyse and evaluate their decisions</i>	Creative and Media	Metacognitive	“5. Evaluation and Checking Goal: to assure coverage and understanding of the text" (Howell and Caros, 2006, p3)
<i>Require the use of innate Psychomotor skills (such as sketching)</i>	Engineering	Psychomotor	“3.0 Conformation. Ability to integrate aptitudes and perform acts with ascribed qualities and characteristics to the point of skills recognition" (Hauenstein, 1998, p98)

Results

The frequency of poles coded with a category was calculated to indicate the goodness of fit between experts' views and established taxonomies. Table 5 gives the frequency and percentage of poles coded as corresponding (or not corresponding) to a taxonomy. There were 722 poles in total.

Table 5: Frequency and percentage of the poles assigned to each taxonomy

	Frequency of poles	Percentage of the total poles
Affective taxonomy	65	9%
Cognitive taxonomy	224	31%
Interpersonal taxonomy	66	9%
Metacognitive taxonomy	61	8%
Psychomotor taxonomy	80	11%
Coded with one or more taxonomies	382	53%
Not coded with any taxonomy	340	47%

The data had some correspondence with all the taxonomies. The majority of the poles (53 per cent) were coded as corresponding with one or more taxonomies. The Cognitive taxonomy corresponded with far more poles than any of the other taxonomies.

Additionally, 47 per cent of the poles did not relate to demands as defined by the taxonomies. These poles related to *specification features* such as the certification outcomes (grades, versus competent and not yet competent) or the mode of assessment (such as a written examination versus portfolio). For more details see [REDACTED]

More fine grained analyses were also conducted to see how well the taxonomies covered each subject and qualification. These showed that each taxonomy was used to code one or more poles from each type of qualification, and each of the subjects.

Discussion

The aim of this research was to develop a re-usable framework for comparing the specification demands of cognate units from different types of qualifications. The strategy for the research was to compare experts' views of demands with established taxonomies. It was found that a reasonable section of the interview data corresponded to the established taxonomies. Therefore, the taxonomies were deemed suitable to use as the basis of a framework for comparing specification demands.

Many educational taxonomies exist. Their main purposes are classifying educational objectives and curriculum and test development. These taxonomies were originally developed using academic subjects and vocational education. However, only a minority of comparability studies explicitly consider these taxonomies. The present study differed in that it related experts' views of demands to established taxonomies, as well as ascertaining the relationship between taxonomies, subjects and types of qualification. The number of experts per subject in the present study was not large and therefore the study shared one of the limitations of previous studies. However, the present research resulted in a potentially re-usable framework by considering the experts' views plus established taxonomies. This was unlike comparability studies that resulted in subject specific scales.

Experts' views of demands were coded as corresponding with established taxonomies. This indicated that there was some correspondence between experts' views of demands and established taxonomies. Pole(s) from each subject corresponded with each taxonomy. This meant that all the taxonomies had some relevance to each subject.

The findings helped the researchers to further clarify the concept of specification demands (the level of knowledge, skills and understanding required of typical learners to successfully complete the course leading to an award or certificate). The findings showed that these demands might be in the Affective, Cognitive, Interpersonal, Metacognitive and Psychomotor domains and they might be explicit or implicit in the specification. Specification demands are intentional inputs; they are not about outputs or how the typical learners actually perform. Specification demands are about the whole unit; they are not about individual assessment tasks, nor the assessment mode itself such as an examination question paper. Typical learners are the majority of the members of the target group, not the most or least able learners or subgroups.

Experts expressed some poles in the context of their professional practice. Arguably this reflected that the experts' knowledge of demands was *situated knowledge*. In other words the experts had years of specialist knowledge and sophisticated mental models of their area. These mental models were influenced by the social context of educational and vocational practice including the physical structure, purposes of practice and other actors. For further details about situated cognition and related issues, see Collins and Gentner (1995), Hennessy (1993) and Schraw (2006). Where experts articulated demands which were situated it could be hard for researchers to disentangle the demands from the context. This could be one of the reasons why Pollitt et al. (2007) found that researchers have varied levels of success in asking experts to articulate demands.

Although it was found that all the taxonomies were relevant to the interview data, the Cognitive taxonomy corresponded with more poles than any of the other taxonomies. This could well reflect experts' experience of curriculum and assessment. It has been claimed that curriculum and assessment tend to be dominated by the cognitive domain (Bloom et al., 1956). This may be because cognitive skills may be easier to test than some of the other domains. Moreover, it has been suggested that there are cognitive aspects to all the domains (Krathwohl et al., 1964; Moseley, Baumfield, Higgins, Lin, Miller, Newton, Robson, Elliott, and Gregson, 2004; Simpson, 1966). This cognitive focus was contrary to the ideal view that vocational education, general education and vocational practice aim to develop individual's knowledge and skills in several domains (Adeyemo, 2010; Hauenstein, 1998; Hsu and Hsieh, 2009; Knutson, 2003; Krathwohl et al., 1964; Moseley et al., 2004; Ormerod, 1983; Simpson, 1966; Timma, 2005).

Comparability studies generally have not considered the relationship between taxonomies, subjects and qualification types. This article attempted to address this gap in the literature. It was found that most experts from three subjects articulated views of demands that corresponded to all of the established taxonomies. Additionally, analysis of the interview data illustrated that each taxonomy corresponded with some data associated with each type of qualification. This meant that generally all the taxonomies were relevant to each subject and each qualification type.

Conclusion

The main research findings were that:

- Experts could articulate demands in the Affective, Cognitive, Interpersonal, Metacognitive and Psychomotor domains. Additionally, the demands were embedded within the context of the educational or vocational practice of their expertise.
- A reasonable proportion of the experts' views of demands corresponded with established taxonomies which were based on robust research (Hauenstein, 1998; Howell and Caros, 2006; Rackham and Morgan, 1977) and this finding implied that some of their views of demands were well-founded.
- Some views of demands expressed by the experts and associated with each subject and qualification type corresponded with the taxonomies. Therefore, the taxonomies were relevant to the subjects and qualification types in the current research.

Given the findings, the five taxonomies could be used to form a potentially re-usable framework for comparing the demands of cognate units from different types of qualifications in future comparability research. The approach provided a strong foundation for further research in the complex area of demands.

Postscript

Since this research was carried out, the taxonomies were adapted to construct a research instrument for comparing specification demand. The tool was piloted by [REDACTED] and used successfully by [REDACTED] in a comparability study.

References

- Adams, R., & Pinot de Moira, A. (2000). A comparability study in GCSE French including parts of the Scottish Standard Grade Examination. A review of question paper demand, a cross moderation study and a statistical analysis of results by awarding body. A study based on the Summer 1999 examination and organised by WJEC and AQA on behalf of the Joint Forum for GCSE and GCE. In P. Newton, Baird, J.-A., Goldstein, H., Patrick, H. & Tymms, P. (Ed.), *Techniques for monitoring the comparability of examination standards*. London: Qualifications and Curriculum Authority.
- Adey, P., & Shayer, M. (1994). *Really raising standards. Cognitive intervention and academic achievement*. London: Routledge.
- Adeyemo, S. A. (2010). The need for skill development/acquisition in science, technology and mathematics education (STEME) in Nigeria. *Journal of Science and Technology Education Research, 1*, 1-9.
- 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*. New York: Longman.
- Arlett, S. (2002). A Comparability Study in VCE Health and Social Care, Units 1, 2 and 5: A review of the examination requirements and a report on the cross moderation exercise. A study based on the Summer 2001 examination and organised by AQA on behalf of the Joint Council for General Qualifications. In P. Newton, Baird, J.-A., Goldstein, H., Patrick, H. & Tymms, P. (Ed.), *Techniques for monitoring the comparability of examination standards*. London: Qualifications and Curriculum Authority.
- Arlett, S. (2003). A Comparability Study in VCE Health and Social Care, Units 3, 4 and 6: A review of the examination requirements and a report on the cross moderation exercise. A study based on the Summer 2002 examination and organised by AQA on behalf of the Joint Council for General Qualifications. In P. Newton, J.-A. Baird, H. Goldstein, H. Patrick & P. Tymms (Eds.), *Techniques for monitoring the comparability of examination standards*. London: Qualifications and Curriculum Authority.
- Bannister, D. (1965). The Rationale and Clinical Relevance of Repertory Grid Technique. *British Journal of Psychiatry, 111*, 977-982.
- Berger, J. L., Kipfer, N., & Buchel, F. (2008). Effects of Metacognitive Intervention in Low-Performing Vocational Students. *Journal of Cognitive Education and Psychology, 7*(3), 337-367.
- 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*. London: Longmans.
- Cohen, L., & Manion, L. (1994). *Research Methods in Education* (4th ed.). London and New York: Routledge.
- Coles, M., & Matthews, A. (1995). Fitness for purpose: a means of comparing qualifications. A report to Sir Ron Dearing. In P. Newton, J.-A. Baird, H. Goldstein, H. Patrick & P. Tymms (Eds.), *Techniques for monitoring the comparability of examination standards*. London: Qualifications and Curriculum Authority.
- Coles, M., & Matthews, A. (1998). Comparing qualifications – Fitness for purpose. Methodology paper. In P. Newton, J.-A. Baird, H. Goldstein, H. Patrick & P. Tymms (Eds.), *Techniques for monitoring the comparability of examination standards*. London: Qualifications and Curriculum Authority.
- Collins, A., & Gentner, D. (1995). How people construct mental models. In D. Holland & N. Quinn (Eds.), *Cultural Models in Language and Thought* (pp. 243-266). Cambridge: Cambridge University Press.
- Edwards, E., & Adams, R. (2002). A Comparability Study in GCE AS Geography including parts of the Scottish Higher Grade Examinations: A review of the examination requirements and a report on the cross moderation exercise: A study based on the Summer 2001 Examination and organised by WJEC on behalf of the Joint Council for General Qualifications. In P. Newton, Baird, J.-A., Goldstein, H., Patrick, H. & Tymms, P. (Ed.), *Techniques for monitoring the comparability of examination standards*. London: Qualifications and Curriculum Authority.
- Edwards, E., & Adams, R. (2003). A Comparability Study in GCE Advanced Level Geography (Including the Scottish Advanced Higher Grade Examinations): A review of the examination requirements and a report on the cross moderation exercise. In P. Newton, Baird, J.-A., Goldstein, H., Patrick,

- H., & Tymms, P. (Ed.), *Techniques for monitoring the comparability of examination standards*. London: Qualifications and Curriculum Authority.
- Ertl, H., & Stasz, C. (2010). Employing an 'employer-led' design? An evaluation of the development of Diplomas. *Journal of Education and Work*, 23(4), 301-317.
- Fearnley, A. (2000). A comparability study in GCSE mathematics. A study based on the summer 1998 examination. Organised by the Assessment and Qualifications Alliance (Northern Examinations and Assessment Board) on behalf of the Joint Forum for the GCSE and GCE. In P. Newton, Baird, J.-A., Goldstein, H., Patrick, H. & Tymms, P. (Ed.), *Techniques for monitoring the comparability of examination standards*. London: Qualifications and Curriculum Authority.
- Fransella, F., Bell, R., & Bannister, D. (2004). *A manual for repertory grid technique*. Chichester: John Wiley and Sons.
- Garrett, J. S. (2010). *The effect of expertise and cognitive demand on temporal awareness in real-time scheduling*. (Master of Science.), Wright State University, Dayton, Ohio, USA. Retrieved from http://rave.ohiolink.edu/etdc/view?acc_num=wright1274824659 (1)
- Gray, E. (2000). A comparability study in GCSE science 1998. A study based on the summer 1998 examination. Organised by Oxford Cambridge and RSA Examinations (Midland Examining Group) on behalf of the Joint Forum for the GCSE and GCE. In P. Newton, Baird, J.-A., Goldstein, H., Patrick, H. & Tymms, P. (Ed.), *Techniques for monitoring the comparability of examination standards*. London: Qualifications and Curriculum Authority.



Education. *International Journal for Cross-Disciplinary Subjects in Education, Special Issue*, 3(3), 1533-1541.

Guthrie, K. (2003). A Comparability Study in GCE Business Studies, Units 4, 5 and 6 VCE Business, Units 4, 5 and 6. A review of the examination requirements and a report on the cross-moderation exercise. A study based on the Summer 2002 examination and organised by EdExcel on behalf of the Joint Council for General Qualifications. In P. Newton, Baird, J.-A., Goldstein, H., Patrick, H. & Tymms, P. (Ed.), *Techniques for monitoring the comparability of examination standards*. London: Qualifications and Curriculum Authority.

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

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

Hennessy, S. (1993). Situated Cognition and Cognitive Apprenticeship: Implications for Classroom Learning. *Studies in Science Education*, 22, 1-41.

- Howell, K., & Caros, J. (2006). Taxonomy of Meta Cognitive Activities: Advanced/Strategic Reading. <http://web.archive.org/web/20100602181922/http://www.wce.wvu.edu/Depts/SPED/Forms/Howell%20-Taxonomy%20of%20Strategic%20Reading.pdf>
- Hsu, L. L., & Hsieh, S. I. (2009). Testing of a measurement model for baccalaureate, nursing students' self-evaluation of core competencies. *Journal of Advanced Nursing*, 65, 2454-2463.
- Hughes, S., Pollitt, A., & Ahmed, A. (1998). *The development of a tool for gauging the demands of GCSE and A Level exam questions*. Paper presented at the British Educational Research Association, The Queen's University, Belfast. <http://www.cambridgeassessment.org.uk/images/109649-the-development-of-a-tool-for-gauging-the-demands-of-gcse-and-a-level-exam-questions.pdf>
- Isaacs, T. (2010). Educational assessment in England. *Assessment in Education: Principles, Policy & Practice*, 17(3), 315-334.
- Jankowicz, D. (2004). *The Easy Guide To Repertory Grids*. Chichester: John Wiley and Sons.
- Johnson, S., & Cohen, L. (1983). *Investigating grade comparability through cross-moderation*. London: Schools Council.
- Knutson, S. (2003). Experiential Learning in second-language classrooms'. *TESL Canada Journal*, 20, 52-63.
- Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of Educational Objectives, Handbook II: Affective domain*. New York: David McKay.
- 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.
- Marope, P. T. M., Chakroun, B., & Holmes, K. P. (2015). Unleashing the Potential. Transforming Technical and Vocational Education and Training *Education on the move*. Paris, France.
- Moseley, D., Baumfield, V., Higgins, S., Lin, M., Miller, J., Newton, D., Robson, S., Elliott, J., & Gregson, M. (2004). Thinking skill frameworks for post-16 learners: an evaluation. London.
- Munby, H., Zanibbi, M., Poth, C., Hutchinson, N. L., Chin, P., & Thornton, A. (2007). Enhancing workplace learning for adolescents: the use of metacognitive instruction. *Education + Training*, 49(1), 8-24. doi: <http://dx.doi.org/10.1108/00400910710729848>
- Newton, P., Baird, J., Goldstein, H., Patrick, H., & Tymms, P. (2007). *Techniques for monitoring the comparability of examination standards*. London: Qualifications and Curriculum Authority.
- Office of Qualifications and Examinations Regulation. (2011). Performance Descriptions for Foundation, Advanced and Higher Principal Learning Qualifications in Hair and Beauty Studies. Coventry and Belfast.
- Office of Qualifications and Examinations Regulation. (2012). International Comparisons in Senior Secondary Assessment. Full Report. Coventry and Belfast.
- Ormerod, M. B. (1983). A Model to Exhibit the Interdependence of the Cognitive and Affective Domains of Objectives for Use in Science and Technical Teacher Training. *Research in Science and Technological Education*, 1, 119-128.
- Pollitt, A., Ahmed, A., & Crisp, V. (2007). The demands on examination syllabuses and question papers. In P. Newton, J. Baird, H. Goldstein & H. Patrick (Eds.), *Techniques for monitoring the comparability of examination standards* (pp. 166-206). London: Qualifications and Curriculum Authority.
- Pollitt, A., Hughes, A., Ahmed, A., Fisher-Hoch, H., & Bramley, T. (1998). The effects of structure on the demands in GCSE and A level questions. Report to Qualifications and Curriculum Authority.
- Pritchard, J., Jani, A., & Monani, S. (2000). A comparability study in GCSE English Syllabus review and cross moderation exercise. A study based on the Summer 1998 examination. Organised by Edexcel on behalf of the Joint Council for General Qualifications. In P. Newton, Baird, J.-A., Goldstein, H., Patrick, H. & Tymms, P. (Ed.), *Techniques for monitoring the comparability of examination standards*. London: Qualifications and Curriculum Authority.
- Qualifications and Curriculum Authority. (2007a). GCSE (short course) citizenship studies comparability study. London.
- Qualifications and Curriculum Authority. (2007b). GCSE French comparability study. London.
- Qualifications and Curriculum Authority. (2008). Inter-subject comparability studies. London.
- Qualifications and Curriculum Development Authority. (2010). QCDA Tests and exams support. Retrieved 15 October, 2015, from <http://webarchive.nationalarchives.gov.uk/20100202100514/testsandexams.qcda.gov.uk/15862.aspx#S>
- Rackham, N., & Morgan, T. (1977). *Behaviour Analysis in Training*. Maidenhead: McGraw-Hill.

- Schraw, G. (2006). Knowledge: Structure and processes. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 245-264). New Jersey: Lawrence Erlbaum Associates, Inc. .
- Shayer, M. (1978). The analysis of science curricula for Piagetian level of demand. *Studies in Science Education*, 5, 115-130.
- Simpson, E. J. (1966). The classification of educational objectives, psychomotor domain. Illinois: University of Illinois.
- Suto, W. M. I., & Nádás, R. (2009). Why are some GCSE examination questions harder to mark accurately than others? Using Kelly's Repertory Grid technique to identify relevant question features. *Research Papers in Education*, 24(3), 335-377.
- Timma, H. (2005). *Performative Assessment and Worker Identity(ies)*. Paper presented at the National Vocational Education and Research Conference Wodonga. <http://hdl.voced.edu.au/10707/135466>
- Willmott, A. S. (1980). *Twelve years of examinations research: ETRU 1965–1977*. London: Schools Council.

Cambridge Assessment: © University of Cambridge Local Examinations Syndicate 2015