Linking instructional verbs from assessment criteria to mode of assessment

Jackie Greatorex, Jo Ireland, Prerna Carroll and Sylvia Vitello

Paper presented at the Journal for Vocational Educational and Training conference, Worcester College, Oxford, UK, 3 – 5 July 2015

Research Division

Cambridge Assessment

1 Regent Street

CB2 1GG

Greatorex.j@cambridgeassessment.org.uk

Cambridge Assessment is the brand name of the University of Cambridge Local Examinations Syndicate, a department of the University of Cambridge. Cambridge Assessment is a non-for-profit organisation.

Abstract

The aim of the research is to develop guidance such that qualification developers can select an instructional verb, identify corresponding modes of assessment with established validity and reliability, and use similar pairings of instructional verb and assessment mode in the future. There appears to be no such guidance for summative assessment in school level qualifications in England. The method used in the project had several stages. First, examples of assessments for which research provides evidence of reliability and validity are identified. Secondly, the instructional verb from every third assessment criterion (or similar) is sampled. Thirdly, the associated mode(s) of assessment are noted as well as whether the assessment is internal or external. The data is analysed to identify the mode(s) of assessment associated with each instructional verb. The findings form the guidance mentioned above. Using the guidance does not guarantee valid and reliable assessment because further factors must be considered in qualification development such as regulatory guidance, manageability, context, occupation and cost.

Introduction

Instructional verbs, such as “design” and “recognise”, occur in learning outcomes (or similar). It is good practice to align the learning outcomes with assessment tasks and this is accomplished when they both use the same verbs and require the student to enact the same verbs (Biggs and Tang, 2007). Universities offer guidance on their use; for example, University College Dublin recommends that “the learner recognises” may be assessed using multiple choice tests or essays and “the learner designs” may be assessed using portfolio assessment (Jennings, 2012). There appears to be no similar guidance on current practice for summative assessment in school level vocational qualifications in England. To facilitate the development of such guidance, this research matched instructional verbs from assessment criteria to mode(s) of assessment.

Definitions

Before reviewing literature in the area, it is important to clarify terms used in this paper.

“Assessment criteria” is used here as a catch-all phrase to mean a statement of what the candidate must know or be able to do. The assessment criteria may be assessment objectives, national occupational standards, learning outcomes, grading criteria, goals, knowledge requirements, performance criteria and so forth. “Instructional verb” is used to mean the verb referring to the action that the candidate must perform. For instance, if the assessment criterion were “the learner agrees a plan of action with the client”, then the instructional verb is “agrees”. It is beyond the scope of this research to review terminology in this area but there is a literature with definitions; see, for example, Scott (2011) McMahon (2006) and Chinien and Boutin (2012).

Below we define several “modes of assessment” used in the summative assessment of school level vocational qualifications in England (AlphaPlus Consultancy Limited, 2014, Greatorex, 2005, Harth and Hemker, 2011, Johnson et al., 2013, Ofqual, 2012, Stone and Dearing, 2009).

A “computer based simulation” is when a computer mimics a situation and the candidate interacts with the computer as if they are in the real situation. The candidate’s performance

in the simulation is assessed. An example is a simulation of word processing and spreadsheet software which the candidate uses as if they were using software in a normal environment (Stone and Dearing, 2009).

In “multiple choice tests” the candidate chooses one of several given responses to a question. Generally multiple choice questions are administered on computer and marked automatically.

An “observation” is when the assessor observes the candidate working in their normal everyday work environment and judges which criteria are met (Greatorex, 2005). The assessor makes detailed records of the activities and the criteria/requirements which are met.

“Oral questioning” is when the assessor asks the candidate questions to elicit information and judges whether the candidate’s responses meet the criteria (Harth and Hemker, 2011).

“Portfolio assessment” refers to the candidate compiling a portfolio of evidence to show that they meet the criteria. For instance, assessors’ assessment decisions and feedback may be recorded in a portfolio. Evidence may also include (copies of) the candidate’s work. The portfolio must indicate how the evidence links to the criteria (Harth and Hemker, 2011). Elsewhere, portfolio assessment is referred to as “evidence accumulation” (Cambridge Assessment, 2009).

A “practical assessment” is when the candidate undertakes a hands-on or work-related task, but not as part of the candidate’s job. For instance, the candidate might cook a dish in a kitchen that meets industry standards but not as part of their job. The assessor judges whether the candidate’s work meets the criteria.

A “professional discussion” is a conversation between the assessor and the candidate. The conversation normally focuses on a particular situation or project (Greatorex, 2005). The assessor records the questions and answers. The assessor judges whether the criteria are met.

“Storyboarding” (reflective account/personal statement) refers to a candidate’s account of their work and how that work meets the criteria (Greatorex, 2005). The record may include an explanation of why particular processes were used (Greatorex, 2005) or a reflection on a particular case or job (Harth and Hemker, 2011). Statements are supplemented by additional evidence, such as a piece of work by the candidate, to underpin the claims (Greatorex, 2005).

A “witness testimony” is produced when a client or colleague observes the candidate working in their normal environment and records the candidate’s achievements. The witness judges whether the criteria are met. Witnesses are not qualified assessors and relationships between the witness and candidate must be declared (Greatorex, 2005).

A “work product” is when the candidate provides an example of their work and the assessor judges whether the work meets the criteria (Harth and Hemker, 2011). The work product could be anything from an email to a report to a piece of carpentry.

“Written tests” with open questions require the candidate to read the question and construct a response (answer). An example is a question requiring an essay as an answer. However, constructed responses can also be much shorter than essays.

Each mode of assessment can be designed to be used as an external or internal assessment. “External assessment” is when a person or group of people independent of the candidate sets the assessment criteria and judges the candidate’s performance. In many cases these people are representatives from an Awarding Organisation. There are also cases when an organisation representing an industrial sector sets the national occupational standards (NOS): statements of the knowledge, skills and standards of performance required in a job. For instance, Habia represents the Hairdressing industry and develops NOS (Habia, 2008). When an assessment is automatically marked, for example a multiple choice test, then it is classified as an external assessment.

“Internal assessment” is characterised by the candidate knowing the person (or group) who chooses the assessment mode or task and/or assesses performance.

Taxonomies, frameworks and assessment writing guides

Perhaps the most famous guidance concerning the link between instructional verbs and assessment is Bloom’s taxonomy of educational objectives (Bloom, 1956). Educational objectives have two parts, a noun phrase identifying who is to perform the action and a verb phrase describing the required action (Munzenmaier and Rubin, 2013). For example, the noun phrase might be “the learner” and the verb phrase might be “predicts that the more older siblings a person has the easier they find it to make friends” [1]. Bloom’s taxonomy includes:

- cognitive categories, such as knowledge and comprehension
- verbs associated with each cognitive category.

The verbs within each category show what students may do to illustrate they meet particular objectives (Munzenmaier and Rubin, 2013). The taxonomy is used to classify educational objectives in terms of the cognitive category. For example, teachers and assessment experts may use the taxonomy to judge the cognitive category of the activities underlying grading criteria and goals (Näsström, 2009).

Bloom’s taxonomy may also be used to design assessment tasks that align with cognitive category and the verb in the educational objectives. The relationships between verbs, cognitive categories and tasks are illustrated by Bloom’s wheel (Figure 1). Munzenmaier and Rubin (2013) explain how to use the wheel. The wheel is organised as three concentric shapes. The inner shape contains the cognitive category from Bloom’s taxonomy. The middle shape contains instructional verbs. The outside shape contains a list of activities. For example, to assess learners’ comprehension, the assessment setter creates an educational objective with a verb from the corresponding section of the middle shape and an assessment task from the outer shape. For instance, to assess whether learners’ comprehension regarding information about two variables, a setter may use the learning objective:

[1] Siblings who are born later within a family are generally advantaged in terms of outgoingness, getting along with others, popularity and ease in making friends (Steelman, 1985)

“The learner predicts that the more older siblings a person has the easier they find it to make friends.”

This may be tested by students drawing a graph of the relationship within the data and predicting how easy people find it to make friends from the number of older siblings they have.

The wheel is used in several fields, such as the use of technology in the Army, and technology in schools (Shadrick et al., 2005, Christensen et al., 2006). However, there is no clear connection between the tasks listed in the outer shape and modes of assessment.

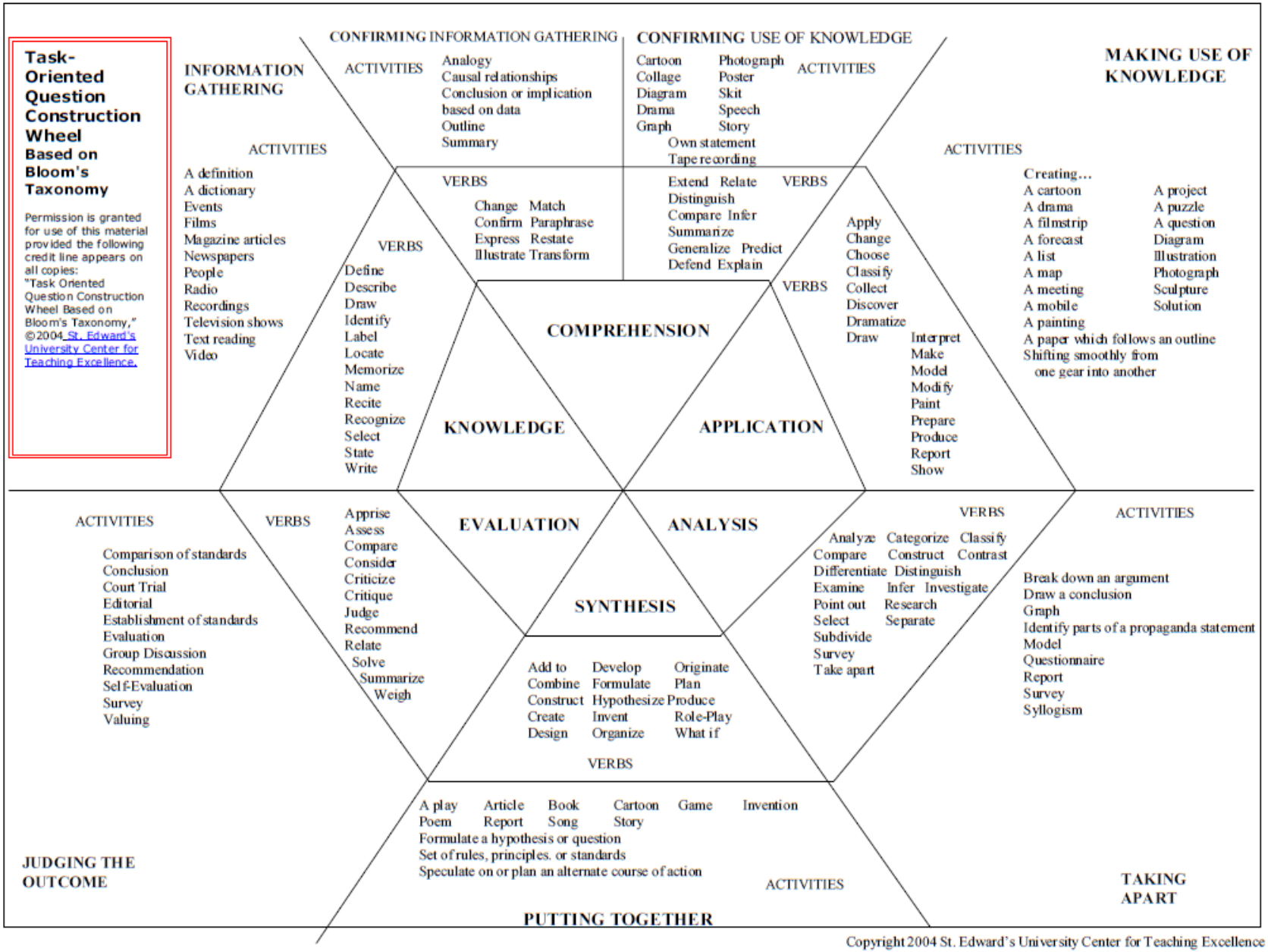
Since Bloom’s taxonomy, many further taxonomies and theory-based frameworks (ways of systematically classifying thinking skills) have been developed. For a comprehensive and insightful review of thinking-skills frameworks and taxonomies, including Bloom’s taxonomy and Gardner’s theory of multiple intelligences, for learners over the age of 16 years, see Moseley et al. (2004). Several of the taxonomies and frameworks they reviewed go beyond thinking skills to the affective domain and conation[2]. Arguably, these taxonomies and frameworks are particularly suitable for use with school level vocational qualifications which include knowledge, skills and understanding from five domains (the affective, cognitive, interpersonal, metacognitive and psychomotor domains) (Greatorex and Shiell, 2012). However, searching related literature did not explain how the frameworks or related verbs link to modes of assessment.

Some universities offer guidance that aligns the actions required of the student with assessment methods. An example is Jennings (2012), which in turn draws from Nightingale et al. (1996). These authors explain that assessment criteria can be divided into types of learning that are aligned with verbs and assessment tasks. Examples from Jennings (2012) which relate to the modes of assessment are described. According to the framework, “demonstrating knowledge and understanding” requires students, for example, to “recall”, “describe”, “identify” or “recognise”, for which suitable modes of assessment include a written examination or multiple choice test. Additionally, “designing, creating, performing” requires students to “design”, “create”, “perform” or “produce”, for which suitable modes of assessment include portfolio assessment. This work suggests links between instructional verbs and a proportion of the modes of assessment.

In short, there appears to be no guidance on aligning instructional verbs with all the modes of assessment used in summative assessment in school level vocational qualifications in England. To facilitate the development of such guidance, this research matched instructional verbs from assessment criteria to mode(s) of assessment.

[2] The part of cognition related to striving, including desire and choosing.

.**Figure 1 Task Orientated Question Construction Wheel Based on Bloom's taxonomy**



Copyright 2004 St. Edward's University Center for Teaching Excellence

Method

Sample

The method had several stages:

1. find examples of assessments for which research provides evidence of reliability and validity
2. sample the instructional verb from every third assessment criterion
3. note the associated mode of assessment e.g. portfolio, multiple choice test
4. note whether the mode of assessment is internal or external

Each of these stages is explained in more detail below.

First, assessments for which there was evidence of validity or reliability were identified (AlphaPlus Consultancy Limited, 2014, Bennett et al., 2003, Bennett et al., 2007, Greatorex, 2005, Harth and Hemker, 2011, Johnson et al., 2013, Newhouse and Njiru, 2009, Ofqual, 2012, Stone and Dearing, 2009, Winther and Koltz, 2013). The choice of references aimed to:

- focus on modes of assessment used in school level vocational qualifications in England
- include a wide range of assessment modes, such as professional discussion, computer based simulation and so on.

Due to the scarcity of literature on the validity and reliability of each mode of assessment, a section of the literature used here was from overseas. Bennett et al. (2003) and Bennett et al. (2007) are about an assessment from the USA and Winther and Koltz (2013) researched an assessment used in Germany.

An assessment was generally a unit in a qualification. There were other examples that were a section of an assessment. For example, Winther and Koltz (2013) provided evidence of the reliability of the business process section of an examination question paper with open responses for higher ability candidates.

Statements of the assessment criteria were sourced generally from specifications and course handbooks. For example, Greatorex (2005) provided evidence of reliability of assessment judgements for the following assessment criteria (from a qualification for National Vocational Qualification Assessors):

“Agree fair, safe, valid and reliable assessment methods”

“Use the past experience and achievements of candidates as part of the assessment of their current competence”

The instructional verb from every third[3] assessment criterion was noted. Using the above example, the instructional verbs were “agree” and “use”. A note was then made of the mode(s) of assessment associated with the assessment criteria, for which there was validity or reliability evidence. Continuing with the above example, the modes of assessment included professional discussion, observation and a personal statement (Greatorex, 2005).

[3] Every third gave a reasonable amount of data.

When the assessment criteria were unavailable a summary of the intended construct was compiled from the literature (Bennett et al., 2003, Bennett et al., 2007, Newhouse and Njiru, 2009, Winther and Koltz, 2013, Stone and Dearing, 2009). For example, problem solving with technology is a mixture of scientific inquiry and computer skills (Bennett et al., 2003). Scientific inquiry is finding information, judging which information is relevant, planning and conducting experiments, monitoring your work, organising and interpreting results and communicating a coherent interpretation (Bennett et al., 2003). Computer skills include: carrying out the mechanical operations of using a computer to fulfil the above, running simulated experiments, gaining information from dynamic visual displays, constructing visual displays of data, sorting data and entering text (Bennett et al., 2003). This yielded several instructional verbs such as “find”, “judge” and “plan”. All the instructional verbs were noted along with the mode of assessment used to test the criteria, for example a computer based simulation.

Finally, the assessment was classified as internally or externally assessed based on the above definitions.

The process outlined above generated a database containing:

- the reference for validity or reliability evidence
- the instructional verb
- the mode of assessment used
- whether it was an internal or external assessment

The assessments were from a variety of occupational areas:

- Assessing and Verifying
- Customer Service
- Electro Technical
- Hairdressing and Barbering
- ICT
- Industrial Management
- Problem Solving with Technology
- Pharmaceutical Science
- Professional Cookery
- Technical Graphics

Analysis

The aim of the analysis was to ascertain the valid or reliable mode(s) of assessment used with each instructional verb. The data were searched to identify the mode(s) of assessment used with each instructional verb.

Findings

Assessment developers can use the following table to select an instructional verb, identify corresponding mode(s) of assessment with established validity and reliability using internal assessment (indicated by "I"), or external assessment (designated by "E") or both (signified by "●"), and then use similar pairings in the future, thus, applying a principle of evidence based practice. However, this does not guarantee valid and reliable assessment because additional factors such as context must be taken into account.

Table 1 Instructional verbs and the associated modes of assessment

Instructional Verbs	Professional discussion	Observation	Storyboarding	Work product	Oral questioning	Witness testimony	Computer based multiple choice	Computer based Simulation	Portfolio	Written test with open responses	Practical
Adapt		I			I				I		
Advise	I	I	I		I				I		
Agree	I	I	I	I	I	I			I	I	
Analyse	I			I	I	I			I	I	I
Apply		I			I				I		●
Assemble											I
Calculate											I
Carry out								E			
Check	I	I	I	I	I	I			I	I	I
Collect	I	I	I								
Comb		I							I		
Communicate								E			
Compare										●	
Complete		I							I		
Comply	I	I	I	I	I	I			I	I	
Conduct	I	I	I	I	I	I		E	I	I	
Confirm	I	I	I	I	I	I			I	I	
Construct								E			
Cook											I
Define											I
Demonstrate										I	I
Describe							E			●	I
Detail							E				
Determine											I
Discuss										I	I
Ensure	I	I	I	I	I	I			I	I	
Enter								E			
Establish		I			I				I		

Instructional Verbs	Professional discussion	Observation	Storyboarding	Work product	Oral questioning	Witness testimony	Computer based multiple choice	Computer based Simulation	Portfolio	Written test with open responses	Practical
Evaluate	I			I	I	I			I	I	I
Execute										E	
Explain	I	I	I	I	I	I	E		I	●	I
Feedback	I	I	I								
Find								E			
Finish											I
Fix	I	I	I	I	I	I			I	I	
Follow	I	I	I	I	I	I			I	I	
Get								E			
Give	I	I	I								
Identify	I	I	I	I	I	I	E		I	I	I
Indicate											I
Inform	I	I	I	I	I	I			I	I	
Interpret								E			
Judge								E			
Know	I	I	I	I	I	I			I	●	
List							E			E	I
Make											I
Monitor								E		E	
Obtain	I	I	I	I	I	I			I	I	
Organise								E			
Outline							E			I	I
Plan								E		E	
Position		I			I				I		
Prepare											I
Produce											●
Protect	I	I	I	I	I	I			I	I	
Provide	I	I	I	I	I	I			I	I	
Recognise							E				
Recommend		I			I				I		
Record							E				
Remove		I			I				I		
Replenish		I							I		
Resolve	I	I	I								
Review	I	I	I	I	I	I			I	I	
Run								E			
Solve								E			

Instructional Verbs	Professional discussion	Observation	Storyboarding	Work product	Oral questioning	Witness testimony	Computer based multiple choice	Computer based Simulation	Portfolio	Written test with open responses	Practical
Sort								E			
State							E			E	I
Store											I
Summarise										I	
Understand	I	I	I	I	I	I			I	●	
Use	I	I	I	I	I	I		E	I	I	I
Number of words aligned with mode of assessment	25	32	23	20	27	20	9	16	30	31	24

The research aimed to ascertain the link between instructional verbs and modes of assessment in the particular case of summative assessment in school level vocational qualifications.

The results show that the external assessments were computer based multiple choice test, computer based simulation, written test and practical assignment. These assessment modes were also used in internal assessments with the exception of computer based multiple choice. The remaining modes of assessment occurred only as internal assessments, for example observation.

Many instructional verbs corresponded only with internal assessments, for example “fix”. About a third of the instructional verbs which matched to external assessment, such as “identify”, also corresponded with internal assessment. The instructional verbs which corresponded only with external assessment, for example “interpret”, mostly related to a computer based simulation designed to assess computer skills as well as the product and process of scientific inquiry.

The final row of Table 1 indicates the number of verbs that were aligned with each mode of assessment. Portfolio and observation were both aligned with a high proportion of instructional verbs and generally the same instructional verbs. This accords with previous research reporting that a great deal of assessment in vocational qualifications is undertaken as observations, and the outcomes of observations are generally part of the evidence in portfolios (Harth and Hemker, 2011). It suggests that observation is a key approach to assessment which can be validly and reliably used to assess a broad range of skills and knowledge.

Portfolios often contain records of professional discussions, observations, work products, oral questioning and witness testimonies (Greatorex, 2005, Harth and Hemker, 2011). Many of the instructional verbs align with portfolio assessment and several of the modes of

assessment found in portfolios, as would be expected. However, a striking feature of Table 1 is that some verbs, such as “give” and “collect” do not align with portfolio assessment but do align with parts of portfolios like observation. This is because there was evidence for reliability of judgements for modes of assessment like observation but judgements at the portfolio level were not researched in studies such as Greatorex (2005).

Another mode of assessment which was aligned to many instructional verbs was a written test. This suggests that written tests can be used to assess a variety of knowledge and skills, or that written tests are often used as they are a familiar mode of assessment. Ferns and Moore (2012) also found that written tests were frequently used to assess fieldwork (work placement) in many nonmedical courses in two Australian higher education institutions.

Finally, the multiple choice tests were aligned with the fewest instructional verbs, perhaps because several fields in higher education note that multiple choice questions may be a valid and reliable test of knowledge, but not other learning domains (Bashook, 2005, Albino et al., 2008, McDonald, 2013, Lievens, 2012, Ferns and Moore, 2012). In contrast, several fields in higher education deem observation fit for assessing knowledge and cognition as well as interpersonal, affective and physical skills (Ferns and Moore, 2012, Winckel et al., 1994, Beard, 2008).

One of the starting points for an assessment developer are regulations the assessment must meet, such as national occupational standards or Ofqual criteria (Ofqual, 2011). The regulations often contain assessment criteria for qualifications. In the case of hairdressing the national occupational standards stipulate that:

“You must be able to comb out your client’s hair in a way suitable for achieving the desired look, when necessary” (SkillsActive, 2014, page 2).

From this extract of national occupational standards an assessment developer could then judge that the instructional verbs are “be able to” and “comb”. Table 1 shows that there is research evidence that “comb” has been validly or reliably assessed using observation (as part of portfolio assessment) in an internal assessment context. This confirms that using this mode of assessment is likely to be good practice. However, the assessment developer must also consider several other factors. For instance, any regulations or guidance about assessment mode must be adhered to such as the Department for Education (2014) requirement that vocational qualifications must contain a certain amount of external assessment to be in school performance tables. Furthermore, the cost of the mode of assessment and internal or external assessment arrangements must be considered.

We compared our findings with one guide to developing assessments in universities (Jennings, 2012). Of the modes of assessment, multiple choice test, written examinations and portfolios were found in the guide. Multiple choice tests and written examinations were aligned with the instructional verbs “describe” and “identify”, in both our research and the university guide. Additionally, multiple choice tests were aligned with the instructional verb “recognise” in both our research and the university guide. This verifies our research findings. There were, nevertheless, some differences. For example, in the university guide “produce” was aligned with portfolio assessment whereas in our research it was aligned with practical assessments. This difference may be due to several factors including the context of the instructional verb and that our findings do not give an exhaustive list of the modes of

assessment which may be aligned with instructional verbs. Nonetheless, Table 1 is one of many sources of information (detailed above) that can contribute to qualification development.

Regarding future directions a further layer of analysis may be to classify the assessment criteria into learning domains and then to link the learning domains to certain verbs and modes of assessment. Such an analysis, and resulting guide for assessment developers, would reflect the structure of other guidance such as Jennings (2012) and Nightingale et al. (1996). Another area of further research would be to add the results of further validity and reliability research to the database and reanalyse the data, thus updating the guidance.

In summary, the results enable assessment developers to select an instructional verb, identify corresponding mode(s) of assessment with established validity and reliability, and then use similar pairings of instructional verb and assessment mode(s) in the future, thereby building on good practice. Using the findings, however, does not guarantee valid and reliable assessment because additional considerations such as context need to be accounted for. Furthermore, the research findings need to be considered alongside other factors such as regularity requirements, cost, and manageability, which must be incorporated in assessment development.

References

ALBINO, J. E. N., YOUNG, S. K., NEUMANN, L. M., KRAMER, G. A., ANDRIEU, S. C., HENSON, L., HORN, B. & HENDRICSON, W. D. 2008. Assessing Dental Students' Competence: Best Practice Recommendations in the Performance Assessment Literature and Investigation of Current Practices in Predoctoral Dental Education *Journal of Dental Education,* 72**,** 1405-1435.

ALPHAPLUS CONSULTANCY LIMITED. 2014. Validation of vocational qualifications. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/371838/2014-02-11-validation-of-vocational-qualifications-final-report.pdf [Accessed February 2014].

BASHOOK, P. G. 2005. Best practices for assessing competence and performance of the behavioural health workforce. *Administration and policy in Mental Health,* 32**,** 563-592.

BEARD, J. D. 2008. Assessment of surgical skills of trainees in the UK. *Annals of the Royal College of Surgeons in England,* 90**,** 282-285.

BENNETT, R. E., JENKINS, F., PERSKY, H. & WEISS, A. 2003. Assessing complex problem solving performances. *Research Memorandum* [Online]. Available: http://www.ets.org/research/policy_research_reports/publications/report/2003/imdu [Accessed 8 June 2015].

BENNETT, R. E., PERSKY, H., WEISS, A. R. & JENKINS, F. 2007. Problem Solving in Technology-Rich Environments. *Research and Development Series* [Online]. Available: http://nces.ed.gov/nationsreportcard/pdf/studies/2007466.pdf [Accessed 3 June 2015].

BIGGS, J. B. & TANG, C. 2007. *Teaching for Quality Learning at University,* Maidenhead, Open University Press/McGraw Hill.

BLOOM, B. S. (ed.) 1956. *Taxonomy of educational objectives: the classification of educational goals. Handbook 1, Cognitive domain,* London: Longman.

CAMBRIDGE ASSESSMENT. 2009. The Cambridge Approach. Principles for designing, administering and evaluating assessment. Available: http://www.cambridgeassessment.org.uk/ca/digitalAssets/188934_cambridge_approach.pdf [Accessed 12th September 2012].

CHINIEN, C. & BOUTIN, F. 2012. Informing enhanced standard formats for the Red Seal Program. Available: http://www.red-seal.ca/images/Informing_enhanced_standard_Final_Report_EN.pdf.
CHRISTENSEN, R., OVERALL, T. & KNEZEK, G. 2006. Personal Educational Tools (PETs) for Type II Learning. *Computers in the Schools: Interdisciplinary Journal of Practice, Theory, and Applied Research,,* 23**,** 173-189.
DEPARTMENT FOR EDUCATION. 2014. Vocational Qualifications for 16 to 19 year olds 2017 and 2018 performance tables : technical guidance for awarding organisations Annex F : 2018 Addendum added. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/412526/16-19_qualifications_technical_guide_2017_and_2018_performance_tables.pdf [Accessed 19 June 2015].
FERNS, S. & MOORE, K. 2012. Assessing student outcomes in fieldwork placements: An overview of current practice. *Asia-Pacific Journal of Cooperative Education,* 13**,** 207-224.
GREATOREX, J. 2005. Assessing the evidence: Different types of NVQ evidence and their impact on reliability and fairness. *Journal of Vocational Education and Training,* 57**,** 149-164.
GREATOREX, J. & SHIELL, H. 2012. Piloting a method for comparing the demand of vocational qualifications with general qualifications. *Research Matters: A Cambridge Assessment Publication***,** 29-36.
HABIA 2008. SKAG21 Provide hairdressing consultation services. *National Occupational Standards.* Habia.
HARTH, H. & HEMKER, B. 2011. On the reliability of results in vocational assessment: The case of work-based certifications. *Ofqual Reliability Programme* [Online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/377920/2011-03-16-on-the-reliability-of-results-in-vocational-assessment.pdf [Accessed 27 August 2015].
JENNINGS, D. 2012. The design of multiple choice questions for assessment. *UCD Teaching and learning/Resources.* Dublin: University College Dublin.
JOHNSON, S., JOHNSON, R., MILLER, L. & BOYLE, A. 2013. Reliability of vocational assessment: an evaluation of level 3 electro-technical qualifications. *Ofqual Reliability Programme* [Online]. Available: http://dera.ioe.ac.uk/17676/1/2013-01-17-c-and-g-reliability-of-vocational-assessment.pdf [Accessed 3 June 2015].
LIEVENS, F. S., PAUL R. 2012. The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *Journal of Applied Psychology,* 97**,** 460-468.
MCDONALD, M. E. 2013. *The Nurse Educator's Guide to assessing learning outcomes,* United States of America, Jones and Bartlett Learning, LLC, an Ascend Company.
MCMAHON, T. 2006. Teaching for more effective learning: Seven maxims for practice. *Radiography,* 12**,** 34-44.
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.
MUNZENMAIER, C. & RUBIN, N. 2013. Bloom's taxonomy: What's old is new again. *Perspectives* [Online]. Available: http://educationalelearningresources.yolasite.com/resources/guildresearch_blooms2013%20(1).pdf [Accessed 27 August 2015].
NÄSSTRÖM, G. 2009. Interpretation of standards with Bloom's revised taxonomy: a comparison of teachers and assessment experts. *International Journal of Research and Method in Education,* 32**,** 39-51.
NEWHOUSE, C. P. & NJIRU, J. N. 2009. Using digital technologies and contemporary psychometrics in the assessment of performance on complex practical tasks. *Technology, Pedagogy and Education,* 18**,** 221-234.

NIGHTINGALE, P., TE WIATA, I. T., TOOHEY, S., RYAN, G., HUGHES, C. & MAGIN, D. 1996. Assessing Learning in Universities Professional Development Centre. Available: http://jisctechdis.ac.uk/assets/documents/resources/database/id235_assessing_learning_in_universities.pdf [Accessed 12 June 2015].

OFQUAL 2011. Criteria for the Diploma Qualifications in Society, Health and Development at Foundation, Higher and Advanced Levels. Coventry.

OFQUAL. 2012. Findings from the Review of Level 3 NVQ Diploma in Hairdressing Qualifications. Available: http://dera.ioe.ac.uk/id/eprint/14008.

SCOTT, I. 2011. The Learning Outcome in Higher Education: Time to think again? *Worcester Journal of Learning and Teaching*.

SHADRICK, S. B., LUSSIER, J. W. & HINKLE, R. 2005. Concept development for future domains: A new method of knowledge elicitation. *Technical Report.* United States Army Research Institute for the Behavioural and Social Sciences.

SKILLSACTIVE 2014. SKAAH4 Cut natural hair using basic techniques. London: National Occupational Standards.

STEELMAN, L. C. & POWELL, B. 1985. The social and academic consequences of birth order: real, artifactual, or both? *Journal of Marriage and the Family,* 47**,** 117-124.

STONE, A. & DEARING, M. 2009. Simulations: a case study of City & Guilds' newest assessment. *Education & Training,* 51**,** 422-433.

WINCKEL, C., P. , REZNICK, R. K., COHEN, R. & TAYLOR, B. 1994. Reliability and construct validity of a Structured Technical Skills Assessment Form. *The American Journal of Surgery,* 167**,** 423-427.

WINTHER, P. D. H. E. & KOLTZ, V. K. 2013. Measurement of vocational competences: an analysis of the structure and reliability of current assessment practices in economic domains. *Empirical Research in Vocational Education & Training,* 5.

Cambridge Assessment: © University of Cambridge Local Examinations Syndicate 2015