Received 25 October 2017
Revised 19 January 2018
Accepted 29 January 2018

Benchmarking higher education programs through alignment analysis based on the revised Bloom's taxonomy

Kwok Hung Lau
RMIT University, Melbourne, Australia
Tri Khai Lam
RMIT University Vietnam, Ho Chi Minh City, Vietnam
Booi Hon Kam
College of Business, RMIT University, Melbourne, Australia
Mathews Nkhoma
RMIT University Vietnam, Ho Chi Minh City, Vietnam, and
Joan Richardson
College of Business, RMIT University, Melbourne, Australia

Abstract

Purpose – The purpose of this paper is to propose a scalable quantitative approach to evaluate alignment within and between courses and programs in higher education for benchmarking purpose.

Design/methodology/approach – The revised Bloom's taxonomy, which combines a cognitive process dimension and a knowledge dimension, is used as a basis for categorizing national standards, program and course learning outcomes (CLOs) and assessment methods. Alignments between programs and national standards, programs and courses and assessment tasks and courses are then measured using a series of Cohen's κ statistics. Two undergraduate business programs offered at an Australian university were used as examples to demonstrate the proposed method as an alignment evaluation tool.

Findings – The findings reveal that the two sample programs are better aligned with national standards than with their respective constituent courses. The degree of alignment between CLOs and assessment methods varies from course to course within the programs. This might be related to the lack of clarity of some learning outcome statements and the complexity of certain assessment methods.

Research limitations/implications – This study lends insight into the use of an alignment mapping for benchmarking academic programs in higher education. To serve mainly as an illustration of the proposed approach, the case study is limited to two undergraduate business programs offered at the same university.

Practical implications – Universities can use the proposed approach to benchmark their academic programs against the national standards and similar programs offered by other competing educational institutions. The alignment indices can also serve as yardsticks to continuously improve the consistencies within and among academic programs to ensure quality.

Originality/value – The proposed method offers a consistent basis to compare the degrees of alignment of different higher education programs with national standards and their respective constituent courses, hence enabling benchmarking for continuous improvement. It also reveals how the alignment between different parameters in teaching and learning can be improved, thereby facilitating incremental learning and enhancing student performance.

Keywords Alignment, Higher education, Benchmarking, Academic programs, Cohen's κ statistic, Revised Bloom's taxonomy

Paper type Research paper

Benchmarking: An International Journal
Vol. 25 No. 8, 2018
pp. 2828-2849
© Emerald Publishing Limited
1463-5771
DOI 10.1108/BIJ-10-2017-0286

1. Introduction

Same as in the business sector, benchmarking plays an important role in higher education. By comparing its administrative processes and instructional models against those of other institutions, a university can improve its performance and strengthen its competitive position.

Benchmarking also helps overcome internal resistance to change, provide a structure for external evaluation, promote communications among universities and encourage information sharing in the industry (Gunasekaran, 2002). The implementation of benchmarking has been considered as a competing tool for excellence in higher education institutions (Tasopoulou and Tsiotras, 2017). Successful benchmarking requires a structured approach (Asif, 2015). In this regard, alignment analysis is a commonly used method for benchmarking in education (Biggs, 1996; Webb, 1997b; Sukanjanaporn *et al.*, 2015). When applied at the strategic level, alignment refers to how well policy elements in a system work together to guide instruction and student learning (Webb, 1997b). At the operational level, it refers to how well courses and programs are associated to provide a comprehensive curriculum based on national or trade standards (Biggs, 1996). Benchmarking through alignment analysis at operational level can help identify areas for improvement in curriculum design, thereby enhancing the performance and competitiveness of an education institution in the long run. This also meets the needs of the industry for aligning competencies of business graduates with the industry requirements (Azevedo *et al.*, 2012; Bowker, 2017).

Alignment between different elements of teaching and learning implies a match between learning goals and educational pedagogies (Webb, 1997b), signifying that curriculum and assessment are in agreement with each other or different parts of the curriculum are appropriately linked together (Klimoski and Amos, 2012). Studies conducted in the past generally support the concept of aligning assessment with learning outcomes (Anderson, 2002; Biggs, 1996; Case *et al.*, 2004; Knight, 2002; Näsström and Henriksson, 2008; Nicol and Macfarlane-Dick, 2006). For example, constructive alignment – a principle advocated by Biggs (1996) and used for devising teaching and learning activities as well as assessment tasks – emphasizes that alignment between all aspects of the curriculum helps ensure the achievement of the desired learning objectives (Borrego and Cutler, 2010). Biggs' (1996) framework assumes that alignment can lead to deeper and more meaningful learning and to real improvements in teaching. For example, curriculum and instruction can be aligned to facilitate incremental learning from lower to higher cognitive levels as the course progresses. Correspondingly, assessment tasks can be designed at different cognitive levels to tie in with the depth and breadth of learning undertaken by students. This is in line with Cohen's (1987) idea of "instructional alignment," a concept underpinned by two prevailing learning theories: cognitive constructivism (Piaget, 1968) and social constructivism (Vygotsky, 1978). Given the importance of alignment, a systematic way to measure it will be essential to the enhancement of curriculum design and assessment methods of a higher education program in the long run (Meyers and Nulty, 2008). The measurement can also facilitate comparison of alignment between programs within an educational institution to ensure internal consistency, or across different educational institutions to set benchmark for continuous improvement.

There were many endeavors to develop alignment measures in education to attain the designed goals (Ananda, 2001; Rothman *et al.*, 2002; Webb, 1997a). Webb (1997b) proposed the criteria for judging alignment based on five general categories: content focus, articulation across grades and ages, equity and fairness, pedagogical implication and system applicability. Porter (2002) suggested measuring alignment between the content of instruction and the content of instructional material using an alignment index. While various descriptive or qualitative approaches to curriculum alignment have been proposed (see Crespo *et al.*, 2010; Davies, 2000; Nusche, 2008; Scott, 2011), numerical or quantitative methods to evaluate alignment between program learning outcomes (PLOs) and course learning outcomes (CLOs), or between CLOs and assessment tasks, are relatively few. Also, most alignment studies in teaching and learning used single course or a particular field of study as the subject of investigation (see Biggs, 1996; Davies, 2000; Webb, 1997a, b; Witzig *et al.*, 2014). There is comparatively little research on alignment of an entire program with national requirements and standards or between learning outcomes of a program and those of its constituent courses. To facilitate objective comparison between programs for benchmarking, hence improvement in curriculum design, a numerical measure to systematically assess alignment between learning objectives and assessment methods of the constituent courses in a program as well as between PLOs and national standards is needed. In this study, the following research question is raised:

RQ1. How can curriculum alignment in higher education at program level, both internally and externally, be measured in a relatively simple and consistent manner?

To answer the question, we propose the use of the revised Bloom's taxonomy (RBT) (Anderson *et al.*, 2001; Krathwohl, 2002) and Cohen's κ statistic (Cohen, 1960, 1968) to develop a scalable method to measure alignment within and between higher education programs.

2. Literature review

2.1 Reasons for alignment in teaching and learning

Researchers generally agree that the effective teaching and learning relates to close alignment of an education program in terms of learning outcomes, teaching and learning approaches, assessment techniques and course evaluation methods (Biggs, 1996; Chadwick, 2004; Valsraj and Lygo-Baker, 2006). The need for alignment is underpinned by prevailing learning theories, such as cognitive constructivism (Piaget, 1968) and social constructivism (Vygotsky, 1978). The cognitive-constructivist framework requires instructors to create environments in which they, and their students, are encouraged to think and explore so as to achieve deep understanding. Structuring curriculum around primary concepts and aligning curriculum with assessment are critical dimensions of constructivist pedagogy (Brooks and Brooks, 1999). From a social-constructivist perspective, learning relies on social interaction and collaboration to make the learning materials meaningful. Therefore, a course content presented through lectures should be accompanied by assessment tasks in which learners must reflect on and use the new information they acquired (Redden *et al.*, 2007). Herman and Webb (2007) contend that there are three main alignments that each education system is expected to set in place: alignment of assessment with curriculum standards, alignment of learning outcomes with curriculum standards and alignment of assessment with learning outcomes. In other words, curriculum design, learning objectives and assessment tasks need to be aligned to ensure student learning is in line with the intended learning outcomes. Previous studies have also shown that aligning curriculum with teaching method can contribute to student performance (Baker, 2004; Porter, 2002; Rothman *et al.*, 2002; Webb, 2007). While improvement in student performance is relatively easy to gauge, measuring alignment between curriculum, instruction and assessment is not a straightforward process.

2.2 Alignment approaches and methodologies

Two alignment approaches, namely traditional rating and matching technique, are commonly used in teaching and learning. Traditional rating methods rely on experts in the discipline to rate the degree of alignment between a stated objective and a test item on a multipoint scale (Harris and Brown, 2010; Herman *et al.*, 2005). Matching techniques involve asking experts to choose the objective which they consider aligning best with an item (Case *et al.*, 2004; Webb, 1997b). While the two methods are distinct in approach, the complexity in curriculum alignment often requires a thorough analysis of content and intellectual skills using both traditional rating and matching technique (D'Agostino *et al.*, 2008).

To determine alignment between curriculum, instruction and assessment, analysis of content to enable matching is usually the first step. For example, Koh and Neuman (2009) assessed the content of a family child care course to determine the alignment between the syllabus and the course activities. Similarly, Stayton *et al.* (2012) used content analysis to determine alignment of state standards with national personnel standards in early childhood special education. Likewise, Witzig *et al.* (2014) proposed the use of a 35-item Biotechnology Instrument for Knowledge Elicitation scored on a three-point scale to evaluate the content of a biochemistry course to establish alignment between guiding concepts and course topics.

In measuring curriculum alignment, researchers tend to rate the alignment of curriculum standards based on a list of criteria (Biggs, 1996; Näsström and Henriksson, 2008; Webb, 1997b). The raters then compare the curriculum materials against the set criteria to determine the degree of alignment (Nasser *et al.*, 2014). For example, Porter (2002) used content maps and an alignment index to measure the match between the content of instruction and the instructional materials as experienced by teachers and students. This numerical method has been applied to evaluate alignment between assessment and standards (Polikoff *et al.*, 2011; Porter *et al.*, 2011), alignment of teachers' instruction and assessment (Polikoff, 2012a, b), alignment of teachers' instruction and their contributions to student learning (Polikoff and Porter, 2014) and alignment between curriculum guidelines and examinations (Liang and Yuan, 2008). The recurring usage of the method suggests that there is merit with numerical measure of alignment. In theory, Porter's (2002) alignment index can be used to measure the extent of alignment of anything that can be content analyzed. In practice, the method has been used mainly to compare standards, assessment and teaching in science, mathematics and quantitative subjects (Porter and Smithson, 2002). Despite its comprehensiveness, popularity of the method can be affected by its need to generate a matrix of topic descriptors and cognitive demand categories as well as the level of coverage for each topic (Näsström and Henriksson, 2008).

2.3 Need for program level alignment tool

Educators generally agree that it is important to align curriculum with program objectives as educational experience should encourage, support and reward students for mastering program learning objectives (Allen, 2004; Kift, 2009; Plaza *et al.*, 2007; Uchiyama and Radin, 2009). A well-designed higher education program can develop expertise of graduates across all four domains of learning – cognitive, affective, psychomotor and conative – and equip them to enter an increasingly competitive global environment (Reeves, 2006). Cohesive curriculum design or mapping – a means of showing the links between elements of curriculum and displaying its essential features in a clear and succinct manner – is often employed to systematically align program content with program objectives, which will then be used to develop CLOs, course activities and grading schemes (Huba and Freed, 2000). Program assessment through the use of program alignment mapping tools, such as Program Alignment Matrix developed by Liu *et al.* (2010), can increase understanding of campus-wide effectiveness of education (Allen, 2004). However, most of the available curriculum alignment tools are descriptive in nature and designed to determine alignment of a single course. The literature on learning and teaching lacks numerical alignment measures for an entire program. Although quantitative tools, such as Porter's (2002) alignment index can, in theory, be applied to every course in a program to ascertain the overall alignment, the amount of work involved can be significant. As such, developing a relatively simple and scalable tool for measuring degree of alignment between different elements of a degree program at the program and course levels could supplement the current inadequacy.

2.4 RBT as a basis for categorization

Numerical alignment methods usually require a reference framework to create content maps for comparison. As Porter (2002, p. 3) contented, "a single language for measuring content ensures description at a consistent level of depth and specificity. [T]he language allows alignment to be measured across a large number of instructional materials and instructional practices." In this regard, the taxonomy developed by Bloom *et al.* (1956) (or Bloom's taxonomy, in short), which has been extensively employed as a strategy for creating content in curriculum to impart learning, can be an appropriate candidate. The original Bloom's taxonomy comprises three overlapping domains: cognitive, psychomotor and affective (Pickard, 2007). In the cognitive domain, the taxonomy provides definitions for each of the six major cognitive categories, each broken into sub-categories. They include knowledge, comprehension, application, analysis, synthesis and evaluation. Anderson *et al.* (2001) extended Bloom's taxonomy by affixing a knowledge dimension to it, which comprises factual, conceptual, procedural and metacognitive categories (Krathwohl, 2002). With minor adjustment in sequential order, the six categories in the cognitive process dimension were converted to their active verb counterparts, including remember, understand, apply, analyze, evaluate and create. As a method of classifying educational goals for student performance evaluation, the original taxonomy provides a basis for educators to categorize levels of learning, in terms of expected outcomes for a given education program. The revised taxonomy creates an intersection of knowledge and cognitive process categories to support design of learning strategies, as well as to facilitate learning assessment.

With cognitive process and knowledge as two distinctive dimensions, the RBT can serve as a framework for gauging curriculum alignment, which may include a number of objectives, a variety of instructional activities and different types of assessment tasks, regardless of whether the analysis is based on the curriculum of a single course or an entire program (Anderson, 2002). Using the RBT, measuring learning outcomes could take into consideration various levels of learning. In general, the RBT can be used to classify instructional activities based on the criteria of both the cognitive process and the knowledge transferred. To measure alignment, the classified activities can be translated into codes and an index or statistic that gauges the agreement between two sets of codes can then be applied (Bakeman and Gottman, 1997; Cohen, 1960; Holt *et al.*, 2015).

2.5 κ statistic for measurement of alignment

In addition to a framework, numerical alignment measures also need a systematic approach to calculate a score to gauge the degree of alignment. In this regard, Cohen's (1960) κ statistic is suitable. The statistic has been widely used in the field of content analysis due to a number of advantages over other measures (Carletta, 1996; Jakobsson and Westergren, 2005). A κ statistic (or coefficient) of inter-judge agreement for nominal scale measures the agreement between two judges on their independent assessment of a subject matter. It represents the proportion of agreement beyond chance (Daly and Bourke, 2000) and is considered a measure of "true" agreement (Cohen, 1960). Derivatives of κ have been used for analyzing hypotheses concerning rater disagreement and interpreted as proportionate reduction-in-error measures (von Eye and von Eye, 2005). The application of κ in teaching and learning is also widespread but the statistic is mainly used to measure inter-rater reliability or agreement in alignment studies (see Nathan and Kim, 2009; Braasch and Goldman, 2010; Chapman *et al.*, 2010; Corte *et al.*, 2013; Fyssa and Vlachou, 2015). Cohen's κ requires a contingency table of classifications by two raters to calculate the degree of inter-rater reliability. In this study, this attribute is leveraged to measure the agreement between two grid maps – for example, one for national standards and another for PLOs – created using the RBT. A contingency table can then be compiled based on the number of matched and mismatched grids between the two maps to calculate the statistic representing their level of alignment. In view of its robustness and ease of interpretation, Cohen's κ is considered appropriate for use as an index for measuring agreement between learning outcomes and set standards.

3. Methodology

For the purpose of benchmarking, this study put forward a scalable approach to evaluate alignment between national standards, programs learning outcomes, its constituent courses and their assessment tasks. The proposed method uses the RBT (Anderson *et al.*, 2001; Krathwohl, 2002) as a basis to categorize national standards, s PLOs and CLOs as well as assessment methods. The categorization of each set of learning outcomes is presented on a map with 24 grids, representing the 6×4 matrix of the RBT with cognitive processes as columns and knowledge types as rows. Cohen's kappa statistic (κ) for the agreement between two grid maps – each represents a set of learning outcomes or assessment methods – is then calculated as a measure of the alignment. The procedure is repeated at different levels to enable a systematic and consistent comparison between programs. Figure 1 shows the concept of this approach.

To illustrate the procedure of the proposed approach and demonstrate its value, two undergraduate business programs in different disciplines offered at an Australian university were used as examples for comparison of their respective alignment with the national standards and the constituent discipline courses. As a basis for categorization, the RBT is used to examine the extent to which: the PLOs of the two undergraduate programs align with the national standards; the CLOs of the constituent courses of the programs align with their PLOs; and the assessment methods of each constituent course align with their CLOs. In each case, the degree of alignment is measured using Cohen's κ statistic. The relative alignment of the two programs is then compared using an index generated by taking the average of the κ coefficients.

3.1 Categorization of learning outcomes

In this study, the RBT is used as a basis for categorizing learning outcomes and objectives of assessment methods to determine alignment. As Krathwohl (2002, p. 213) explains, intended learning outcomes are generally written in such a way that comprises (1) a description of action on a subject matter and (2) the subject matter or content itself. They typically consist of (1) a verb, or verb phrase, to represent a cognitive process and (2) a noun, or noun phrase, representing the subject matter or content. The former fits in with the cognitive process dimension and the latter the knowledge dimension of the RBT. For example, a typical CLO may read as follows: "Upon completion of the course, students will be able to understand the key concepts and theories in the field of supply chain management." In this example, the cognitive process action verb is "understand" and the Knowledge type is "concepts and theories." So, if the six categories of the cognitive process in the RBT – remember, understand, apply, analyze, evaluate and create – are labeled as from C1 to C6 and the four knowledge types – factual, conceptual, procedural and metacognitive – as from K1 to K4, the CLO in the example can be categorized as C2K2. In other words, by matching the action verb(s) and the content noun(s) in a CLO against the corresponding verbs and nouns in the RBT, it is possible to categorize the CLO into one or more of the 24 cells in a two-dimensional matrix formed by the six cognitive processes and the four knowledge types. Anderson (2002) also concurs that the RBT table with 24 cells

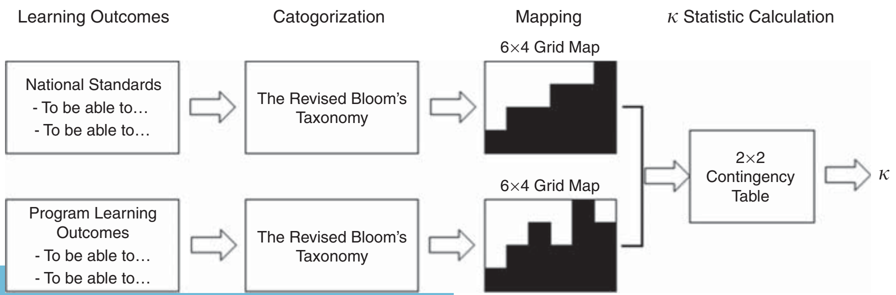


Figure 1. Basic concept of the proposed alignment evaluation approach

enables teachers to use it as a framework to examine and enhance curriculum alignment. The generic nature of the knowledge types in the RBT table also implies that it can be used for diverse subject matters. Using the cognitive verbs and knowledge nouns in the RBT to prepare program and CLOs as well as course assessment tasks, the framework can be used as a standard for categorizing learning and teaching elements at national, program and course level, thus making the proposed approach scalable.

3.2 Measurement of degree of alignment

Cohen's (1960) κ is a commonly used statistic to measure agreement for nominal scale (i.e. data recorded in categories). It was originally intended to assess agreement between two or more observers, with similar skills and experiences in the subject matter, in categorizing items into two classes. The coefficient is popular because it measures the "true" agreement beyond that is expected by chance as shown in the following equation:

$$\kappa = \frac{\text{observed agreement}-\text{chance agreement}}{1-\text{chance agreement}}. \quad (1)$$

For example, the agreement between two observers A and B in their independent classifications of a set of subjects into two classes 1 and 2 can be measured using κ by creating a contingency table summarizing the classification results (Table I).

The proportion of observed agreement (P_o), i.e. where the two independent observers actually give identical classifications can be calculated as follows:

$$P_o = \frac{(a+d)}{n}. \quad (2)$$

The proportion of agreement expected by chance (P_c), i.e. where the two independent observers give the same results entirely due to chance can be calculated as follows:

$$P_c = \frac{(a+b)(a+c)+(b+d)(c+d)}{n^2}. \quad (3)$$

Then, based on Equation (1), κ can be calculated using the following equation:

$$\kappa = \frac{P_o-P_c}{1-P_c}. \quad (4)$$

When $P_o=1$, κ has a maximum value of 1 and the agreement is regarded as perfect. When $P_c=P_o$, κ has a value of 0 indicating that the agreement is no better than chance. When $P_o=0$, κ has a minimum value of $-P_c\,/\,(1-P_c)$ which is negative suggesting a worse than chance agreement (Bartko and Carpenter, 1976).

While Cohen's κ statistic is commonly used to measure reliability of clinician's rating in medical research (see Donner and Klar, 1996; Kraemer *et al.*, 2002; McHugh, 2012), the coefficient has also been widely applied to compare maps for differences in attributes

Table I.
Contingency table used in calculating Cohen's κ statistic

	Observer B's classification		
Observer A's classification	Class 1	Class 2	Total
Class 1	a	b	$a+b$
Class 2	c	d	$c+d$
Total	$a+c$	$b+d$	$n=a+b+c+d$

Notes: a, b, c, d are actual counts; a and d are counts of agreements while b and c are counts of disagreements (see Lau and Kam, 2005; Monserud and Leemans, 1992; Visser and de Nijs, 2006). To measure the strength of agreement using the κ statistic, a scale ranging from poor to almost perfect has been proposed by Landis and Koch (1977) (see Table II).

3.3 Combining the RBT with κ to measure alignment of higher education program

If the two dimensions of the RBT are seen as map references with the horizontal axis representing the cognitive process and the vertical axis the knowledge type, the various learning outcomes and assessment methods with specific cognitive processes and knowledge types can be taken as grids marked on a map. By comparing two maps with different marked grids, a κ statistic can be calculated to measure the degree of their agreement. Using this technique, alignment between national standards and programs, courses and programs and assessment tasks and courses can be measured on a scalable and consistent basis. Following the principle of backward design in education (Childre *et al.*, 2009; Cho and Trent, 2005), the alignment between national standards and a program, then the program and its constituent courses, and finally the courses and the assessment tasks, can be evaluated in sequence and measured using a series of κ statistics. While the individual statistics representing alignment at the different levels – between national standards and programs, courses and programs, and assessment tasks and courses – can help improve curriculum design and development of appropriate course contents and assessment methods, the overall alignment of the program can be signified by averaging the various statistics to facilitate comparison between programs for benchmarking. Table III summarizes the steps involved in the process which are also visually depicted in Figure 2.

Strength of agreement	Poor	Slight	Fair	Moderate	Substantial	Almost perfect
Value of κ	$\leqslant 0$	0.01–0.20	0.21–0.40	0.41–0.60	0.61–0.80	0.81–1.00

Source: Landis and Koch (1977)

Table II. Interpreting the magnitude of Cohen's κ statistic

Step	Activity	Map
1	Categorize the national standards of the relevant degree program and mark them as grids on a map A using the RBT dimensions as grid references	A
2	Categorize the intended program learning outcomes (PLOs) of the higher education program and mark them as grids on another map B using the RBT dimensions as grid references	B
3	For the constituent courses in the program, categorize the intended course learning outcomes (CLOs) and the assessment methods of each individual course and mark them as grids separately on two maps (C and D) using the RBT dimensions as grid references. Consequently, a series of pairs of maps (Cs and Ds) will be created	Cs, Ds
4	Amalgamate the maps (Cs) representing the intended CLOs of all the constituent courses in the program into one map (E)	E
5	Calculate the κ statistic (κ_N) for agreement between Maps A and B to determine the degree of alignment of the intended PLOs of the program with the national standards	A and B
6	Calculate the κ statistic (κ_P) for agreement between Maps B and E to determine the degree of alignment of the intended PLOs of the program with the intended CLOs of the constituent courses	B and E
7	Calculate the κ statistic (κ_i) for agreement between each pair of Maps C and D to determine the degree of alignment of the intended CLOs of a constituent course with its assessment methods. Consequently, a series of κ statistics ($\kappa_1, \kappa_2, \ldots, \kappa_n$) will be generated and the average value (mean κ_C) will be taken	Cs and Ds
8	Calculate an index (mean κ) representing the overall alignment of the program, from the perspectives of the national standards, the entire program and the constituent courses combined together, by taking the average of the three κ statistics (κ_N, κ_P, mean κ_C)	NA

Table III. Steps in using the RBT and Cohen's κ statistic in measuring alignment

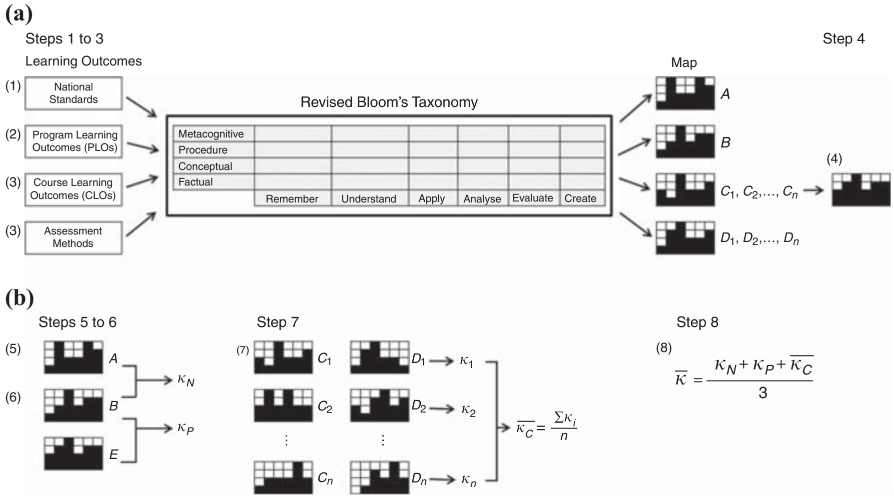


Figure 2. The RBT mapping and the calculation of κ statistics

Notes: (a) Mapping national standards, learning outcomes, and assessment methods onto the revised Bloom's taxonomy (RBT); (b) calculating κ statistics representing degree of alignment between two maps. In the Step 8, where: κ_N, κ statistic at national level; κ_P, κ statistic at program level; $\overline{\kappa_C}$, mean κ statistic at course level

To minimize subjective judgment and bias in the categorization of program and CLOs as well as assessment methods, the categorization process – Steps 1 to 3 in Table III – can be conducted independently by multiple raters using Cohen's (1968) weighted κ to evaluate the agreement between each pair of raters. Any disagreement can then be resolved through elaborated group discussion and further analysis until a consensus is reached. This approach of resolving disagreement is commonly used in alignment studies, such as Porter and Smithson (2001) and Liang and Yuan (2008). In this study, the categorizations of learning outcomes and assessment methods were conducted independently by all authors prior to consolidation. While the majority of the categorizations are identical and therefore taken as final, the few instances of disagreements were resolved through group discussion and deliberation.

4. Illustrative examples

To illustrate the proposed procedures, two undergraduate business programs offered by an Australia university were mapped. One of the two higher education programs is in the discipline of logistics and supply chain management (hereafter referred to as BBus (LSCM)) whereas the other one is in information systems (hereafter referred to as BBus (IS)). Both programs require three years of full-time study and follow the same 8+8+8 program structure – eight 12-credit point courses in each year with four in one semester – as adopted by the university being studied. The first year comprises eight common business courses which are compulsory for all students enrolled in business programs regardless of disciplines. The second and the third year each comprises four discipline courses and four other courses. Students can make use of the eight non-discipline courses (four each in the second and the third year) to take a second major or double minor or a minor with student electives. In this regard, there are over a hundred courses a student can take and, therefore, two students in the same program might have taken two sets of totally different minor or elective courses upon graduation. This design gives students maximum flexibility in pursuing their studies to meet their individual needs. To eliminate bias in comparing the alignment of the two programs, only the eight discipline courses in the respective programs are considered in the calculation of the κ statistics.

The national standards for undergraduate degree programs are specified in the Australian Qualifications Framework (AQF). First introduced in Australia in 1995 as the national policy to underpin the system of qualifications encompassing higher education, vocational education and training and schools, the AQF incorporates the qualifications from each education and training sector into a single comprehensive national qualifications framework. It is a taxonomic structure of levels and qualification types, each of which is defined by a taxonomy of learning outcomes. The AQF levels define the relative complexity and depth of achievement and the autonomy required of graduates to demonstrate that achievement. There are ten levels, with Level 1 having the lowest complexity and Level 10 the highest complexity. The AQF qualification type, a total of 14 from across all education and training sectors, is defined by a descriptor expressed as learning outcomes constructed as a taxonomy of what graduates are expected to know, understand and be able to do as a result of learning. They are expressed in terms of the dimensions of knowledge, skills and the application of knowledge and skills (Australian Qualifications Framework Council, 2013). To a large extent, the knowledge and skills referred to in the AQF correspond to the dimensions of Knowledge and Cognitive Process dimensions in the RBT.

The AQF classifies three-year bachelor degrees as qualification type Level 7 (or AQF Level 7 in short). Table IV summarizes the learning outcome descriptors of this level. The bolded words in the learning outcomes are active verbs (or nouns referring to actions) identified to match against the cognitive processes in the RBT and the italicized words are contents to match against the knowledge types. By matching the AQF Level 7 descriptors against the corresponding RBT's knowledge and cognitive process dimensions (Table V), a map representing the AQF Level 7 standards can be created (Figure 3).

Using the BBus (LSCM) program as an example, the published PLOs and their mapping onto the RBT are shown in Table VI. Again in the PLOs, the bolded words are active verbs identified to match against the cognitive processes and the italicized words are contents to

Headings	Bachelor Degree – AQF Level 7	
		Learning outcomes
Knowledge Graduates of Bachelor Degree will have	K1	A broad and coherent body of *knowledge*
	K2	Depth in the underlying *principles* and *concepts* in one or more disciplines as a basis for independent lifelong learning
Skills Graduates of a Bachelor Degree will have	S1	Cognitive skills to **review**, critically **analyze, consolidate** and **synthesize** *knowledge*
	S2	Cognitive and technical skills to **demonstrate** a broad understanding of *knowledge* with depth in some areas
	S3	Cognitive and creative skills to **exercise** critical thinking and **judgement** in identifying and solving problems with intellectual independence
	S4	**Communication skills** to **present** a clear, coherent and independent exposition of *knowledge* and *ideas*
Application of Knowledge and Skills Graduates of a Bachelor Degree will demonstrate the application of knowledge and skills	A1	**Initiative** and **judgement** in planning, problem solving and decision making in professional practice and/or scholarship
	A2	**Adapt** *knowledge* and skills in diverse contexts
	A3	**Responsibility and accountability** for own learning and professional practice and in **collaboration** with others within broad parameters

Source: Australian Qualifications Framework Council (2013)

Table IV. Learning outcome descriptors of AQF qualification type Level 7 match against the knowledge types in the RBT. The map representing the PLOs created from Table VI is shown in Figure 4.

Similarly, for each of the eight discipline courses comprising the BBus (LSCM) program, the CLOs and the assessment methods can also be mapped onto the RBT, respectively.

Table V. Matching AQF Level 7 descriptors against the revised Bloom's taxonomy (RBT)

AQF Level 7	Revised Bloom's taxonomy (RBT)	
Descriptor	Cognitive process dimension	Knowledge dimension
K1	C1, C2	K1, K2
K2	C2, C3	K1, K2
S1	C4, C5, C6	K1, K2
S2	C3, C4	K2, K3
S3	C5, C6	K4
S4	C2, C3	K1, K2
A1	C5, C6	K3, K4
A2	C5, C6	K2, K3
A3	C5, C6	K4

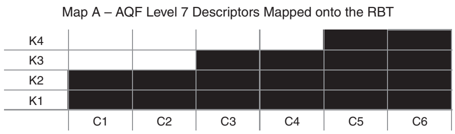


Figure 3. AQF Level 7 descriptors mapped onto the RBT

Table VI. Matching BBus (LSCM) program learning outcomes against the RBT

Program learning outcomes of BBus (LSCM)	Categories in RBT
1. **Utilize** generic business *knowledge* & capabilities to **analyze** business problems, engage reflective practice and **develop** *conceptual* frameworks to inform and improve future practices using appropriate technical tools and language of the field	C3K1, C3K2, C4K1, C4K2, C6K2
2. **Apply** supply chain management *principles* and operational *concepts* to **integrate**, **coordinate** and synchronize supply chain activities to articulate and deliver customer-directed quality outcomes within legal, regulatory, business and ethical frameworks in local and international environments	C3K2, C4K2, C5K2
3. **Develop** operations and integration *processes* to source and **utilize** appropriate technology and applications to support the implementation of logistics *strategies*, automate logistics operations and manage logistics resources to improve supply chain operations with due consideration given to occupational health and safety risks	C3K2, C6K3, C6K4
4. **Leverage** material management knowledge to **implement** *concepts*, *techniques* and applications that underpin material management operations, processes and workflows to allow for efficient procurement, monitoring and control of information and resources	C3K2, C3K3
5. **Use** state-of-the-art distribution practices to **implement** strategic and operational *concepts*, *techniques* and *applications* that underpin distribution and delivery of goods and services for domestic and international markets	C3K2, C3K3
6. **Improvise** creative supply chain design & solutions to research, **plan**, **develop**, **implement** and **evaluate** less conventional supply chain solutions in a dynamic business environment at the *strategic*, tactical and *operational* levels to ensure sustainable business practices	C3K3, C5K3, C6K4
7. **Develop** *interpersonal skills* to engage with others in culturally diverse and technically complex *situations*, to develop *lifelong learning skills* and to become effective project team members	C6K4
8. **Apply** teamwork and *leadership capabilities* to actively lead, engage, influence and work with people of diverse skills and cultural backgrounds within a dynamic business environment to achieve stated business goals and objectives	C3K4

The number of learning outcomes in a course varies from four to eight but the number of assignments is limited to two in addition to the end-of-semester examination. The assignments, which may include online quiz, quantitative analysis exercise, qualitative case study, research essay, investigative project or real-world problem solving depending on whether the lower or higher order learning of students is assessed, can be individual or group work and may comprise a report and a presentation. Taking the first discipline course Introduction to Logistics and Supply Chain Management as an example, the CLOs and the assessment methods with their mapping onto the RBT are shown in Table VII. Once again, the bolded and the italicized words in the learning outcomes and the assessment methods represent the corresponding active verbs (or nouns referring to actions) and knowledge types in the RBT. The maps representing the CLOs and the assessment methods created from Table VII are shown in Figure 5. By combining the maps of the CLOs (i.e. Map C's) of all the eight discipline courses, a map representing the overall learning outcomes of the discipline courses of the BBus (LSCM) program (i.e. Map E) can be created (Figure 6).

Once the maps representing the national standards (AQF Level 7), PLOs, CLOs and the assessment methods in the RBT categories are generated, agreement between a pair of maps can be measured by calculating the corresponding Cohen's κ statistic. For example,

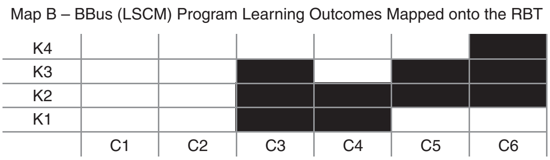


Figure 4. BBus (LSCM) program learning outcomes mapped onto the RBT

Course learning outcomes of introduction to logistics and supply chain management	Categories in RBT
1. **Explain** and **evaluate** key *concepts* and *theories* in the field of logistics and supply chain management	C2K2, C5K2
2. **Interpret** and **apply** the *concepts* of logistics and supply chain management in assisting other functional areas of any business organization	C2K2, C3K2
3. **Build** and **display** appropriate *leadership* and *organizing abilities* in leveraging resources, capabilities, and competencies of a group to critically **analyze** *situations* and **develop** *solutions* to problems	C3K4, C4K3, C6K3
4. **Develop** and **apply** effective *interpersonal skills* and *communication techniques* in working as a team to solve real-world problems in supply chain management	C3K2, C3K3, C6K4
Assessment Methods of Introduction to Logistics and Supply Chain Management	
1. Group Portfolio consisting of a group presentation and a report. Each member in a group is required to select for **investigation** one supply chain *process* that underlines the moving of a product from the suppliers to the end consumers. These *processes* may include demand management, procurement, operations, warehousing, distribution, transportation, *etc*. Student members in each group will have to present their findings on (supply chain *activities* found in the selected *processes*, and the contribution of the activities to supply chain management. The group is also required to write up a formal report, based on the topic selected for the presentation, which fully **analyze** and **illustrate** their selection of the product, the supply chain *processes* and their findings	C2K2, C2K3, C3K4, C4K3, C6K4
2. Individual Online Quiz. Students are required to complete an online quiz that is designed to examine students' **understanding** of logistics and supply chain management *concepts* covered in the lectures and the tutorials	C2K2
3. End-of-Semester Examination. A closed book examination will be held at the end of the semester. The exam is a 2-hour examination consisting of structured and case study questions that examine students' **understanding** and **applications** of logistics and supply chain management *concepts*	C2K2, C3K2

Table VII. Matching learning outcomes and assessment methods of introduction to logistics and supply chain management against the RBT to measure the degree of alignment between the PLOs of the BBus (LSCM) program and the AQF Level 7 descriptors, Maps A and B are overlayed to generate a contingency table and Cohen's κ statistic (κ_N) is calculated using Equations (2)–(4), as shown in Figure 7.

5. Findings and discussion

Using the methodology described in the previous section, a set of Cohen's κ statistics for the BBus (LSCM) and BBus (IS) programs can be calculated and compared. The results are shown in Table VIII. As far as alignment with the national standards is concerned, BBus (IS) has a higher κ statistic (in the moderate category) than BBus (LSCM) (in the Fair category), suggesting that the former is better aligned. When comparing the alignment of PLOs against CLOs, the κ statistics of the two programs both fall in the Fair category, suggesting that both are not well aligned in this aspect. For the alignment of CLOs against assessment methods, an interesting observation can be made. For the BBus (IS) program, the κ statistics of the majority of the Years 1 and 2 courses are lower than those of BBus (LSCM). The situation reverses when the Year 3 courses are compared. Most of the courses in the BBus (IS) programs have a higher degree of alignment than those of the BBus (LSCM) program. Taking all the three levels of alignment – AQF Level 7 vs PLOs, PLOs vs CLOs and CLOs vs assessment methods – into consideration, the BBus (IS) program is better aligned than the BBus (LSCM) program, with the former in the Moderate category, whereas the latter, in the Fair category.

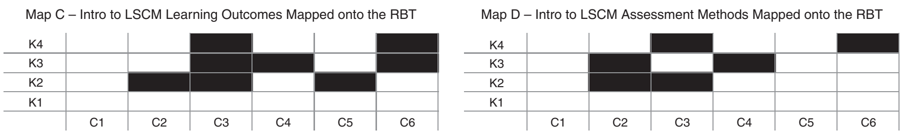


Figure 5. Introduction to logistics and supply chain management course learning outcomes and assessment methods mapped onto the RBT, respectively

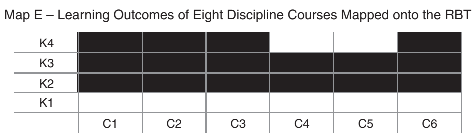
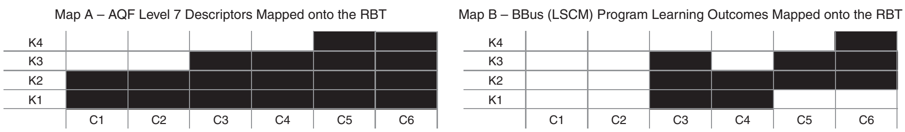


Figure 6. Learning outcomes of all eight discipline courses mapped onto the RBT



		PLOs of BBus (LSCM) (Map B) Shaded	Not Shaded	Row Total
AQF Level 7 Descriptors (Map A)	Shaded	10	8	18
	Not Shaded	0	6	6
	Column Total	10	14	24

Notes: $P_o = (10+6)/24 = 0.667$; $P_c = [(10+8)(10+0)+(8+6)(0+6)]/24^2 = 0.458$; $\kappa = (0.667–0.458)/(1–0.458) = 0.385$

Figure 7. Degree of alignment between AQF Level 7 descriptors and program learning outcomes of BBus (LSCM) measured by Cohen's κ statistic

Alignment	Cohen's κ statistic (κ) BBus (LSCM)	BBus (IS)
(1) AQF Level 7 *vs* PLOs (i.e. Map A vs Map B) – κ_N	*0.385*	*0.500*
(2) PLOs vs All CLOs (i.e. Map B vs Map E) – κ_P	*0.211*	*0.250*
(3) CLOs vs Assessment Methods (i.e. Map C vs Map D) – Discipline Course 1 – Year 1	0.600	0.556
CLOs vs Assessment Methods (i.e. Map C vs Map D) – Discipline Course 2 – Year 2	0.524	0.400
CLOs vs Assessment Methods (i.e. Map C vs Map D) – Discipline Course 3 – Year 2	0.354	0.250
CLOs vs Assessment Methods (*i.e.* Map C vs Map D) – Discipline Course 4 – Year 2	0.500	0.714
CLOs vs Assessment Methods (i.e. Map C vs Map D) – Discipline Course 5 – Year 3	0.294	0.500
CLOs vs Assessment Methods (i.e. Map C vs Map D) – Discipline Course 6 – Year 3	0.474	0.556
CLOs vs Assessment Methods (*i.e.* Map C vs Map D) – Discipline Course 7 – Year 3	0.486	0.556
CLOs vs Assessment Methods (i.e. Map C vs Map D) – Discipline Course 8 – Year 3	0.684	0.556
(4) CLOs vs Assessment Methods (i.e. Map C vs Map D) – Average of (3) – κ_C	*0.489**	*0.511**
(5) Overall for the Program – Average of (1), (2) and (4) – mean κ	*0.362**	*0.420**

Note: *Represents average figure

Table VIII. Comparison of alignment of BBus (LSCM) and BBus (IS)

From a methodological perspective, there are several possible reasons for the observed differences. First, the classification of PLOs and CLOs as well as assessment methods into categories along the two RBT dimensions hinges on the identification of active verbs signifying the cognitive processes and the nouns representing the knowledge types used in the statements. Therefore, whether a PLO or CLO or an assessment method is clearly written using the relevant RBT actions verbs and nouns or not determines to a large extent the success of the classification. For example, the following CLO of a Year 2 discipline course in the BBus (LSCM) program are clearly written using RBT verbs and nouns and can therefore be easily classified:

> Comprehend and evaluate key concepts and theories in the field of transportation and freight logistics.

The RBT active verbs are "comprehend" and "evaluate." They can clearly be grouped under C2 and C5. The RBT nouns are "concepts" and "theories." Again, they can be unambiguously grouped under K2. As a result, the learning outcome can be labeled as C2K2 and C5K2.

In contrast, the following learning outcome of a Year 2 discipline course in the BBus (IS) program is not written in RBT verbs and nouns and therefore cannot be easily classified:

> Manage data in a database system by using data manipulation language (Structured Query Language); SFIA Codes include: PROG Programming/Software Development (Level 2).

The active verb "manage" is broad in meaning and does not augur well with any of the RBT verbs. The closest verb appears to be "apply." Therefore, the cognitive process is classified as C3. The noun "data manipulation language" can be interpreted as knowledge of terminology which is factual knowledge under K1. However, it can also be construed as knowledge of classifications and categories as well as structures, which is conceptual knowledge under K2. Upon extensive deliberation among the independent raters, this learning outcome is labeled as C3K2.

Owing to the way some of the CLOs and the corresponding assessment methods are written, their classifications into the RBT categories can be quite different resulting in significant misalignment. This may account for the low κ statistics observed in the mapping of CLOs against assessment methods. The contrasting examples shown above illustrate the importance of providing clear and comprehensive guidelines when writing CLO statements and assessment tasks. If the RBT is to be adopted as a reference framework for preparing learning outcomes and assessment methods as practiced in many higher education institutes (Heer, 2012; O'Neil and Murphy, 2010), the RBT active verbs and knowledge types should be used to ensure consistency and alignment.

Another possible reason for the variations in alignment at national, program and course levels relates to the characteristics of the reference framework and the disciplines. The RBT is a discipline-free generic taxonomy of cognitive processes and knowledge types. Mapping national standards, such as AQF Level 7, which are also discipline-free, onto the RBT is straightforward. However, PLOs and CLOs are discipline-specific. Although scholars, such as Anderson *et al.* (2001) and Huitt (2011), have proposed to expand the list of synonyms for the six major cognitive processes in the RBT, it is still not possible to encompass all the knowledge acquiring activities, particularly in programs and courses, such as computer programming or information system design, that require special settings or use of complex assessment methods to gauge learning outcomes. This may account for the "fair" to "moderate" alignment of the PLOs with the national standards.

The other possible source of variation in alignment can be attributed to the focus of learning outcome. In categorizing the PLOs and the CLOs, we observed an emphasis on higher-order cognitive processes and knowledge types. Lower-order cognitive process, such as remember (C1) and understand (C2), and factual knowledge type (K1), are relatively less mentioned in the PLO or CLO statements. For example, in the eight PLOs of the BBus (LSCM) program (Table VI), the RBT active verbs of "remember" and "understand" have not been used at all. Factual knowledge type (or generic knowledge) has been mentioned in only one learning outcome. This imbalance is particularly obvious in the CLOs with the majority of them emphasizing the application of concepts, principles and techniques. Consequently, when mapping the CLOs of all the discipline courses of the two example programs onto the RBT and comparing against their respective PLOs, the degrees of alignment in both cases are not high owing to the overemphasis on higher-order knowledge types in the CLOs. Again, the findings highlight the importance of providing clear and comprehensive guidelines for the writing of learning outcome statements and assessment tasks to ensure a balanced and comprehensive capture of knowledge of various types in higher education programs.

Finally, disparity between learning outcomes and assessment design can also be a major source of misalignment. In the two higher education programs being studied, assessment methods of some courses are also not quite aligned with the CLOs. For example, the assessment methods of the e-Supply Chain course in the BBus (LSCM) programs include two group assignments as follows:

> To investigate and to develop a sound understanding on how the latest information technologies maintain its contributions/benefits in e-supply chain coordination.

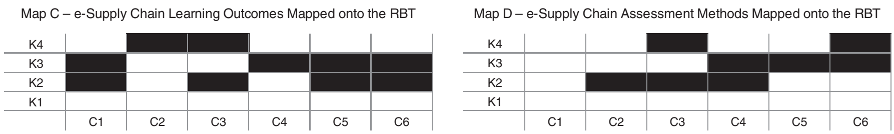
> To investigate and study the current supply chain of an actual company and develop plans to migrate the current supply chain onto the e-platform (e-supply chain).

These two assessment tasks are relatively broad and complex. Furthermore, they both focus on higher-order cognitive processes and knowledge types. The semester-end examination contains an analysis of a case which also emphasizes cognitive process and knowledge of higher-order learning. However, two out of the eight CLOs of the course require the identification of issues and challenges in e-supply chain design and implementation, which falls into the category of lower-order cognitive processes. Consequently, the CLOs and the assessment methods are not well aligned and the resulting κ statistic falls only in the Fair category (see Figure 8). This example illustrates the significance of breaking down a complex assessment method into well-defined simpler tasks so that each of them could align closer with an equally well-defined intended learning outcome.

Table IX summarizes the areas identified in the two illustrative examples that have caused major misalignments and the corresponding improvement measures.

6. Conclusions and implications

In a regulated higher education environment, the quality of a teaching program could be judged by the extent to which its objectives meet the criteria of the nationally prescribed standards. Likewise, the extent to which CLOs could match those of the program within which they sit would be an indicator of quality. Further, the extent to which the assessment tasks set for a course reflecting the intent of the course would be a measure of how well the assessment activities support the achievement of the CLOs. Motivated by studies on the need for benchmarking in higher education (Asif, 2015; Pursglove and Simpson, 2007; Tasopoulou and Tsiotras, 2017; Tee, 2016) and underpinned by previous research on curriculum alignment and assessment framework (Anderson, 2002; Biggs, 1996; Case *et al.*, 2004; Näsström and Henriksson, 2008), this study has developed a systematic and consistent approach to measure alignment between different elements of a program from the perspectives of national standards, program and CLOs as well as assessment tasks to facilitate benchmarking. This is achieved by mapping national standards, PLOs, CLOs and assessment tasks onto the RBT followed by comparing the differences between the respective maps through the use of κ statistics. While previous efforts in measuring curriculum alignment focus primarily on individual courses, the method proposed in this research enables measurement and comparison of alignment at program and course levels, hence supplementing the inadequacy in the literature. Therefore, one of the values of the proposed approach is that it is easy to implement and can be readily applied to measure alignment in teaching and learning at different hierarchies, thus making it scalable and flexible for benchmarking.



		Assessment Methods of e-Supply Chain (Map D)		
		Shaded	Not Shaded	Row Total
CLOs of e-Supply Chain (Map C)	Shaded	5	5	10
	Not Shaded	3	11	14
	Column Total	8	16	24

Notes: P_o = (5+11)/24 = 0.667; P_c = [(5+5)(5+3)+(5+11)(3+11)]/24^2 = 0.528; κ = (0.667–0.528)/(1–0.528) = 0.294

Figure 8. Degree of alignment of e-supply chain course learning outcomes and assessment methods

Causes of Misalignment Identified	Improvement Measures
Learning outcomes and assessment tasks are not written using the standard RBT cognitive action verbs or knowledge types	To provide comprehensive guidelines and examples as references to reduce inconsistency in the writing of learning outcomes and assessment tasks
Current RBT cognitive action verbs and knowledge types are not adequate to describe learning outcomes and assessment tasks of certain disciplines	To enlarge the pool of RBT cognitive action verbs and knowledge types to cover more disciplines
There is a tendency to over-emphasize the use of higher order cognitive processes and knowledge types in writing the learning outcomes and assessment tasks	To provide comprehensive guidelines and examples as references to ensure a balanced coverage of cognitive processes and knowledge types in the writing of learning outcomes and assessment tasks
Assessments tasks can be excessively complex comprising multiple objectives not written in the standard RBT cognitive action verbs or knowledge types	To break down a complex assessment task into smaller and simpler well-defined subtasks written using the standard RBT cognitive action verbs and knowledge types

Table IX. Causes of misalignment and corresponding improvement measures

Another value of the proposed method for measuring alignment is that it facilitates program and curriculum design. Using the RBT framework to categorize learning outcomes and the κ statistic to measure agreement, the method can indicate where the misalignment is and how the alignment can be improved, for example, by adhering to the use of the RBT active verbs and knowledge types in setting the learning outcomes and assessment methods. A low κ statistic reveals room for improvement and points to areas where learning outcomes or learning objectives may need to be revised. The approach can also facilitate the design of assessment methods by breaking up a complex assignment into simpler component tasks to increase their degree of alignment with the learning objectives.

In short, the proposed method can foster a closer alignment between learning outcomes and assessment methods using the RBT as a reference frame. At a higher level, the approach can be applied to increase alignment between degree programs and national standards in curriculum design and assessment methods. At a lower level, the methodology could be employed to reveal gaps in the design of instructional materials and assessment tasks at different cognitive levels and knowledge types. The scalable nature of the approach, grounded on a pedagogically sound and easy to use RBT, makes it an effective tool to enhance alignment of some of the critical factors, such as goals, learner tasks and assessment, to ensure learning success (Reeves, 2006). Continuous improvement in alignment conforms to the premise of cognitive constructivism to promote reflection and deep thinking (Brooks and Brooks, 1999) and agrees with the aim of social constructivism to make learning materials meaningful (Redden *et al.*, 2007). The approach facilitates benchmarking between academic programs within a university or among educational institutions, thereby enabling a better understanding of the university's performance in teaching and learning.

The proposed method is not without limitations. Like other methods in evaluating alignment of learning outcomes, subjective judgment cannot be avoided, particularly in the mapping process when the learning outcomes are not written using RBT active verbs of cognitive processes and nouns of knowledge types. Also, the categorization of learning outcomes is less straightforward when multiple objectives are specified in a single learning outcome or when various tasks are embedded in one assessment method. To overcome this, a group of independent raters can be used to resolve disagreement in matching cognitive processes and knowledge types to help achieve a consensus in categorization (Myford and Wolfe, 2002; Thornton III and Rupp, 2006).

The proposed approach not only provides a tool for measuring alignment and comparing higher education programs across different levels for consistency and benchmarking purposes. It also offers opportunities for research on how learning outcomes and assessment methods could be structured using the RBT to promote incremental learning from lower cognitive processes and knowledge types to higher order ones, for example, through collaborative learning by forming small semi-independent student groups (Collier, 1980). Furthermore, effort could also be invested in enlarging the pool of active verbs for the cognitive processes and knowledge types in the RBT to facilitate categorization of both generic and discipline-specific learning outcomes. As such, the proposed method could be applied in a wider scope to compare alignment between programs in different disciplines.

References

Allen, M.J. (2004), *Accessing Academic Programs in Higher Education*, Anker Publishing Company, Boston, MA.

Ananda, S. (2001), *Rethinking Issues of Alignment Under No Child Left Behind*, WestEd, San Francisco, CA.

Anderson, L.W. (2002), "Curricular alignment: a re-examination", *Theory into Practice*, Vol. 41 No. 4, pp. 255-260.

Anderson, L.W., Krathwohl, D.R., Airasian, P.W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., Raths, J. and Wittrock, M.C. (2001), *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*, Longman, New York, NY.

Asif, M. (2015), "Determining improvement needs in higher education benchmarking", *Benchmarking: An International Journal*, Vol. 22 No. 1, pp. 56-74.

Australian Qualifications Framework Council (2013), "Australian qualifications framework", available at: www.aqf.edu.au/sites/aqf/files/aqf-2nd-edition-january-2013.pdf (accessed October 25, 2017).

Azevedo, A., Apfelthaler, G. and Hurst, D. (2012), "Competency development in business graduates: an industry-driven approach for examining the alignment of undergraduate business education with industry requirements", *The International Journal of Management Education*, Vol. 10 No. 1, pp. 12-28.

Bakeman, R. and Gottman, J.M. (1997), *Observing Interaction: An Introduction to Sequential Analysis*, 2nd ed., Cambridge University Press, New York, NY.

Baker, E.L. (2004), "Aligning curriculum, standards, and assessments: fulfilling the promise of school reform", CSE Report No. 645, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), University of California, Los Angeles, CA.

Bartko, J.J. and Carpenter, W.T. (1976), "On the methods and theory of reliability", *Journal of Nervous and Mental Disease*, Vol. 163 No. 5, pp. 307-317.

Biggs, J. (1996), "Enhancing teaching through constructive alignment", *Higher Education*, Vol. 32 No. 3, pp. 347-364.

Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H. and Krathwohl, D.R. (1956), *Taxonomy of Educational Objectives: The Classification of Educational Goals*, David McKay Company, New York, NY.

Borrego, M. and Cutler, S. (2010), "Constructive alignment of interdisciplinary graduate curriculum in engineering and science: an analysis of successful IGERT proposals", *Journal of Engineering Education*, Vol. 99 No. 4, pp. 355-369.

Bowker, L. (2017), "Aligning accreditation and academic program reviews: a Canadian case study", *Quality Assurance in Education*, Vol. 25 No. 3, pp. 287-302.

Braasch, J.L.G. and Goldman, S.R. (2010), "The role of prior knowledge in learning from analogies in science texts", *Discourse Processes*, Vol. 47 No. 6, pp. 447-479.

Brooks, J.G. and Brooks, M.G. (1999), *In Search of Understanding: The Case for Constructivist Classrooms*, 2nd ed., Association for Supervision and Development, Danvers, MA.

Carletta, J. (1996), "Assessing agreement on classification tasks: the kappa statistic", *Computational Linguistics*, Vol. 22 No. 2, pp. 249-254.

Case, B.J., Jorgensen, M.A. and Zucker, S. (2004), "Alignment in educational assessment, assessment report", Pearson Education, San Antonio, TX, pp. 1-13.

Chadwick, S.M. (2004), "Current products and practices: curriculum development in orthodontic specialist registrar training: can orthodontics achieve constructive alignment?", *Journal of Orthodontics*, Vol. 31 No. 3, pp. 267-274.

Chapman, L., Greenfield, R. and Rinaldi, C. (2010), "'Drawing is a frame of mind': an evaluation of students' perceptions about reading instruction within a response to intervention model", *Literacy Research and Instruction*, Vol. 49 No. 2, pp. 113-128.

Childre, A., Sands, J.R. and Pope, S.T. (2009), "Backward design", *Teaching Exceptional Children*, Vol. 41 No. 5, pp. 6-14.

Cho, J. and Trent, A. (2005), "'Backward' curriculum design and assessment: what goes around comes around, or haven't we seen this before?", *Taboo: The Journal of Culture And Education*, Vol. 9 No. 2, pp. 105-122.

Cohen, J. (1960), "A coefficient of agreement for nominal scales", *Educational and Psychological Measurement*, Vol. 20 No. 1, pp. 37-46.

Cohen, J. (1968), "Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit", *Psychological Bulletin*, Vol. 70 No. 4, pp. 213-220.

Cohen, S.A. (1987), "Instructional alignment: searching for a magic bullet", *Educational Researcher*, Vol. 16 No. 8, pp. 16-20.

Collier, K.G. (1980), "Peer-group learning in higher education: the development of higher order skills", *Studies in Higher Education*, Vol. 5 No. 1, pp. 55-62.

Corte, M.V., Brok, P.D., Kamp, M. and Bergen, T. (2013), "Teacher research in Dutch professional development schools: perceptions of the actual and preferred situation in terms of the context, process and outcomes of research", *European Journal of Teacher Education*, Vol. 36 No. 1, pp. 3-23.

Crespo, R.M., Najjar, J., Dernt, M., Leony, D., Neumann, S., Oberhuemer, P., Totschnig, M., Simon, B., Gutierrez, I. and Delgado Kloo, C. (2010), "Aligning assessment with learning outcomes in outcome-based education", IEEE EDUCON Education Engineering 2010, Madrid, pp. 1239-1246.

D'Agostino, J.V., Welsh, M.E., Cimetta, A.D., Falco, L.D., Smith, S., VanWinkle, W.H. and Powers, S.J. (2008), "The rating and matching item-objective alignment methods", *Applied Measurement in Education*, Vol. 21 No. 1, pp. 1-21.

Daly, L.E. and Bourke, G.J. (2000), *Interpretation and Uses of Medical Statistics*, 5th ed., Blackwell Science, Oxford.

Davies, A. (2000), "Effective assessment in art and design: writing learning outcomes and assessment criteria in art and design, project report", CLTAD, University of the Arts, London.

Donner, A. and Klar, N. (1996), "The statistical analysis of kappa statistics in multiple samples", *Journal of Clinical Epidemiology*, Vol. 49 No. 9, pp. 1053-1058.

Fyssa, A. and Vlachou, A. (2015), "Assessment of quality for inclusive programs in Greek preschool classrooms", *Journal of Early Intervention*, Vol. 37 No. 3, pp. 190-207.

Gunasekaran, A. (2002), "Benchmarking in education", *Benchmarking: An International Journal*, Vol. 9 No. 1, pp. 6-7, available at: https://doi.org/10.1108/bij.2002.13109aaa.001

Harris, L.R. and Brown, G.T.L. (2010), "Mixing interviews and questionnaire methods: practical problems in aligning data", *Practical Assessment, Research & Evaluation*, Vol. 15 No. 1, pp. 1-19.

Heer, R. (2012), "A model of learning objectives–based on a taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives", Center for Excellence in Learning and Teaching, Iowa State University, Ames, IA, available at: www.celt.iastate.edu/wp-content/uploads/2015/09/RevisedBloomsHandout-1.pdf (accessed October 25, 2017).

Herman, J.L. and Webb, N.M. (2007), "Alignment methodologies", *Applied Measurement in Education*, Vol. 20 No. 1, pp. 1-5.

Herman, J.L., Webb, N.M. and Zuniga, S.A. (2005), "Measurement issues in the alignment of standards and assessments", CSE Technical Report No. 653, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), University of California, Los Angeles, CA.

Holt, A.H., Young, C., Keetch, J., Larsen, S. and Mollner, B. (2015), "The greatest learning return on your pedagogical investment: alignment, assessment or in-class instruction?", *PLoS ONE*, Vol. 10 No. 9, pp. 1-19.

Huba, M.E. and Freed, J.E. (2000), *Learner-centered Assessment on College Campuses: Shifting the Focus from Teaching to Learning*, Allyn & Bacon, Needham Heights, MA.

Huitt, W. (2011), "Bloom et al.'s taxonomy of the cognitive domain", Educational Psychology Interactive, Valdosta State University, Valdosta, GA.

Jakobsson, U. and Westergren, A. (2005), "Statistical methods for assessing agreement for ordinal data", *Scandinavian Journal of Caring Sciences*, Vol. 19 No. 4, pp. 427-431.

Kift, S. (2009), "Articulating a transition curriculum to scaffold and to enhance the first year student experience in Australian higher education", final report for ALTC Senior Fellowship Program, Queensland University of Technology, Brisbane, QLD.

Klimoski, R. and Amos, B. (2012), "Practicing evidence-based education in leadership development", *Academy of Management Learning and Education*, Vol. 11 No. 4, pp. 685-702.

Knight, P.T. (2002), "Summative assessment in higher education: practices in disarray", *Studies in Higher Education*, Vol. 27 No. 3, pp. 275-286.

Koh, S. and Neuman, S.B. (2009), "The impact of professional development on family child care: a practice-based approach", *Early Education and Development*, Vol. 20 No. 3, pp. 537-562.

Kraemer, H., Periyakoil, V.S. and Noda, A. (2002), "Kappa coefficients in medical research", *Statistics in Medicine*, Vol. 21 No. 2, pp. 2109-2129.

Krathwohl, D.R. (2002), "A revision of Bloom's taxonomy: an overview", *Theory into Practice*, Vol. 41 No. 4, pp. 212-218.

Landis, J.R. and Koch, G.G. (1977), "The measurement of observer agreement for categorical data", *Biometrics*, Vol. 33 No. 1, pp. 159-174.

Lau, K.H. and Kam, B.H. (2005), "A cellular automata model for urban land use simulation", *Environment and Planning B: Planning and Design*, Vol. 32 No. 2, pp. 247-263.

Liang, L.L. and Yuan, H. (2008), "Examining the alignment of Chinese national physics curriculum guidelines and 12th-grade exit examinations: a case study", *International Journal of Science Education*, Vol. 30 No. 13, pp. 1823-1835.

Liu, M., Wrobbel, D. and Blankson, I. (2010), "Rethinking program assessment through the use of program alignment mapping technique", *Communication Teacher*, Vol. 24 No. 4, pp. 238-246.

McHugh, M.L. (2012), "Interrater reliability: the kappa statistic", *Biochemia Medica*, Vol. 22 No. 3, pp. 276-282.

Meyers, N.M. and Nulty, D.D. (2008), "How to use (five) curriculum design principles to align authentic learning environments, assessment, students' approaches to thinking and learning outcomes", *Assessment & Evaluation in Higher Education*, Vol. 34 No. 5, pp. 565-577.

Monserud, R.A. and Leemans, R. (1992), "Comparing global vegetation maps with the kappa statistic", *Ecological Modelling*, Vol. 62 No. 4, pp. 275-293.

Myford, C. and Wolfe, E.W. (2002), "When raters disagree, then what: examining a third-rating discrepancy resolution procedure and its utility for identifying unusual patterns of ratings", *Journal of Applied Measurement*, Vol. 3 No. 3, pp. 300-324.

Nasser, R., Zaki, E., Allen, N., Al Mula, B., Al Mutawaha, F., Al Bin Ali, H. and Kerr, T. (2014), "Alignment of teacher-developed curricula and national standards in Qatar's national education reform", *International Education Studies*, Vol. 7 No. 10, pp. 14-24.

Näsström, G. and Henriksson, W. (2008), "Alignment of standards and assessment: a theoretical and empirical study of methods for alignment", *Electronic Journal of Research in Educational Psychology*, Vol. 6 No. 3, pp. 667-690.

Nathan, M.J. and Kim, S. (2009), "Regulation of teacher elicitations in the mathematics classroom", *Cognition and Instruction*, Vol. 27 No. 2, pp. 91-120.

Nicol, D. and Macfarlane-Dick, D. (2006), "Formative assessment and self-regulated learning: a model and seven principles of good feedback practice", *Studies in Higher Education*, Vol. 31 No. 2, pp. 199-218.

Nusche, D. (2008), "Assessment of learning outcomes in higher education: a comparative review of selected practices", OECD Education Working Paper No. 15, EDU/WKP(2008)2, Organisation for Economic Co-operation and Development, Paris.

O'Neil, G. and Murphy, F. (2010), "Guide to taxonomy for learning", UCD Teaching and Learning, Dublin.

Piaget, J. (1968), *Six Psychological Studies* (translated by Tenzer, A.), Vintage Books, New York, NY.

Pickard, M.J. (2007), "The new Bloom's taxonomy: an overview for family and consumer sciences", *Journal of Family and Consumer Sciences Education*, Vol. 25 No. 1, pp. 45-55.

Plaza, C.M., Draugalis, J.R., Slack, M.K., Skrepnek, G.H. and Sauer, K.A. (2007), "Curriculum mapping in program assessment and evaluation", *American Journal of Pharmaceutical Education*, Vol. 71 No. 2, pp. 1-8.

Polikoff, M.S. (2012a), "The association of state policy attributes with teachers' instructional alignment", *Educational Evaluation and Policy Analysis*, Vol. 34 No. 3, pp. 278-294.

Polikoff, M.S. (2012b), "Instructional alignment under no child left behind", *American Journal of Education*, Vol. 118 No. 3, pp. 341-368.

Polikoff, M.S. and Porter, A.C. (2014), "Instructional alignment as a measure of teaching quality", *Educational Evaluation and Policy Analysis*, Vol. 36 No. 4, pp. 399-416.

Polikoff, M.S., Porter, A.C. and Smithson, J. (2011), "How well aligned are state assessments of student achievement with state content standards?", *American Educational Research Journal*, Vol. 48 No. 4, pp. 965-995.

Porter, A.C. (2002), "Measuring the content of instruction: uses in research and practice", *Educational Researcher*, Vol. 31 No. 7, pp. 3-14.

Porter, A.C. and Smithson, J.L. (2002), "Alignment of assessments, standards and instruction using curriculum indicator data", paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA, April 1–5.

Porter, A.C. and Smithson, J.L. (2001), "Are content standards being implemented in classroom? A methodology and some tentative answers", in Fuhrmans, S.H. (Ed.), *From the Capitol to the Classroom. Standards-Based Reform in the States*, National Society for the Study of Education, University of Chicago Press, Chicago, IL, pp. 60-80.

Porter, A., McMaken, J., Hwang, J. and Yang, R. (2011), "Common core standards: the new U.S. intended curriculum", *Educational Researcher*, Vol. 40 No. 3, pp. 103-116.

Pursglove, J. and Simpson, M. (2007), "Benchmarking the performance of English universities", *Benchmarking: An International Journal*, Vol. 14 No. 1, pp. 102-122.

Redden, K.C., Simon, R.A. and Aulls, M.W. (2007), "Alignment in constructivist-oriented teacher education: identifying pre-service teacher characteristics and associated learning outcomes", *Teacher Education Quarterly*, Vol. 34 No. 3, pp. 149-164.

Reeves, T.C. (2006), "How do you know they are learning? the importance of alignment in higher education", *International Journal of Learning Technology*, Vol. 2 No. 4, pp. 294-309.

Rothman, R., Slattery, J.B., Vranek, J.L. and Resnick, L.B. (2002), "Benchmarking and alignment of standards and testing", CSE Technical Report No. 566, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), University of California, Los Angeles, CA.

Scott, I. (2011), "The learning outcome in higher education: time to think again?", *Worcester Journal of Learning and Teaching*, No. 5, pp. 1-8.

Stayton, V.D., Smith, B.J., Dietrich, S.L. and Bruder, M.B. (2012), "Comparison of state certification and professional association personnel standards in early childhood special education", *Topics in Early Childhood Special Education*, Vol. 32 No. 1, pp. 24-37.

Sukanjanaporn, P., Sujiva, S. and Pasiphol, S. (2015), "Development of a nursing training programs evaluation: an application of alignment concepts", *Procedia – Social and Behavioral Sciences*, Vol. 191 No. 2, pp. 748-752.

Tasopoulou, K. and Tsiotras, G. (2017), "Benchmarking towards excellence in higher education", *Benchmarking: An International Journal*, Vol. 24 No. 3, pp. 617-634.

Tee, K.F. (2016), "Suitability of performance indicators and benchmarking practices in UK universities", *Benchmarking: An International Journal*, Vol. 23 No. 3, pp. 584-600.

Thornton, G.C. III and Rupp, D.E. (2006), *Assessment Centers in Human Resource Management: Strategies for Prediction, Diagnosis, and Development*, Lawrence Erlbaum Associates, Mahwah. NJ.

Uchiyama, K.P. and Radin, J.L. (2009), "Curriculum mapping in higher education: a vehicle for collaboration", *Innovations in Higher Education*, Vol. 33 No. 4, pp. 271-280.

Valsraj, K.M. and Lygo-Baker, S. (2006), "A balancing act: developing curricula for balanced care within community psychiatry", *Advances in Psychiatric Treatment*, Vol. 12 No. 1, pp. 69-78.

Visser, H. and de Nijs, T. (2006), "The map comparison kit", *Environmental Modelling & Software*, Vol. 21 No. 3, pp. 346-358.

von Eye, A. and von Eye, M. (2005), "Can one use Cohen's kappa to examine disagreement?", *Methodology*, Vol. 1 No. 4, pp. 129-142.

Vygotsky, L. (1978), *Mind in Society*, Harvard University Press, London.

Webb, N.L. (1997a), "Determining alignment of expectations and assessments in mathematics and science education", *NISE Brief*, Vol. 1 No. 2, pp. 1-10.

Webb, N.L. (1997b), "Criteria for alignment of expectations and assessments in mathematics and science education", Research Monograph No. 6, National Institute for Science Education (NISE) Publications, Washington, DC, pp. 1-39.

Webb, N.L. (2007), "Issues related to judging the alignment of curriculum standards and assessments", *Applied Measurement in Education*, Vol. 20 No. 1, pp. 7-25.

Witzig, S.B., Rebello, C.M., Siegel, M.A., Freyermuth, S.K., Izci, K. and McClure, B. (2014), "Building the BIKE: development and testing of the biotechnology instrument for knowledge elicitation (BIKE)", *Research in Science Education*, Vol. 44 No. 5, pp. 675-698.

About the authors

Kwok Hung Lau is Senior Lecturer of Logistics in the School of Business IT and Logistics, Royal Melbourne Institute of Technology University in Australia. He has more than 30 years of working experiences in the private sector, government and the academia that include planning, managerial, research and teaching positions in the field of transport/land use planning and modeling, logistics and supply chain management. His publications appear in various academic journals including *Environment and Planning B: Planning and Design*, *Transaction in GIS*, *International Journal of Production Economics*, *Supply Chain Management: An International Journal* , *International Journal of Physical Distribution and Logistics Management* and *Education + Training*. Kwok Hung Lau is the corresponding author and can be contacted at: charles.lau@rmit.edu.au

Tri Khai Lam is Research Officer, works with undergraduate and postgraduate lecturers and professors. His research interests are divided between teaching and learning in higher and further education and taking a systematic approach to business and management.

Booi Hon Kam is Professor of Logistics and Supply Chain Management at RMIT University. He published widely on supply risk issues and other supply chain related topics. His major areas of research include supply chain management, supply chain risk analysis, information systems in supply chain, port management and higher education.

Mathews Nkhoma is Head at the Department for Business Information Technology and Logistics, RMIT University Vietnam. His major topics of research include zero day attacks and vulnerability research, information systems security, transparency in information security system design, network security investment model, ethical hacking and network defence, network security management, forensic computing and evidence recovery including mobile devices, cybercrime, identity theft, consumer protection, trust and confidence, impact of information and communication technology in education.

Joan Richardson is Deputy Head, Learning and Teaching in the School of Business Information Technology and Logistics, RMIT University. She published widely on IT education issues and received an Australian Learning and Teaching Council citation in 2011 recognizing her exploration of innovative technology-based learning resource delivery. Her major areas of research include information technology on education, information security compliance and governance in IT management.