No: 026

Design of Assessment Instruments for Applied Learning Subjects in Hong Kong

Tak-wah Fu, Hong Kong Examinations and Assessment Authority, Hong Kong
twfu@hkeaa.edu.hk
Kai-yin Bill Kwok, Hong Kong Examinations and Assessment Authority, Hong Kong
kykwok@hkeaa.edu.hk

Abstract

The first Hong Kong Diploma of Secondary Education (HKDSE) Examination will be administered in 2012 to assess the achievement of senior secondary school graduates. Besides four core subjects (Chinese Language, English Language, Mathematics and Liberal Studies), Hong Kong students are free to take two to four elective subjects. These elective subjects include traditional subjects as well as career-oriented and competency-based Applied Learning (ApL) subjects. There are a total of 30 ApL subjects, which will be offered for the first time in September 2010. The diversity of these 30 subjects in terms of context, curriculum design and objectives posts difficult problems to assessment. How will these subjects be graded? How can they be compared? Are the assessment tasks reliable and valid? How will the standards be maintained? All sorts of queries and problems have emerged. Since the results of ApL subjects will be certified in the HKDSE, the Hong Kong Examinations and Assessment Authority (HKEAA) has to face the challenge of devising appropriate assessment instruments to measure the performance of students taking different ApL subjects. The standards of the subjects need to be comparable. A combination of different assessment methods and approaches has to be employed effectively in assessing students' achievement.

Key words: Applied Learning; Competency-based Assessment; Assessment Tasks; Performance Descriptors; Standards

Introduction

Applied Learning is an integral part of the senior secondary curriculum, complementing senior secondary subjects by offering studies with stronger elements of practical learning linked to broad professional and vocational fields. Apart from achieving learning experiences, the introduction of ApL subjects also aims to develop the generic skills that underpin Hong Kong's curriculum framework. The 30 ApL subjects are structured into six areas of studies: (1) Creative Studies, (2) Media and Communication, (3) Business, Management and Law, (4) Services, (5) Applied Science, and (6) Engineering and Production, and are currently run by eleven Course Providers (CPs). In the context of Applied Learning, assessment involves generating and collecting evidence of students' attainment of knowledge, skills, values and attitudes, and judging that evidence against defined standards. Assessment in Applied Learning serves two fundamental purposes: (1) facilitating learning and monitoring student progress, and (2) providing data and information for certification. In order to ensure the consistency of assessment standards, Hong Kong Examinations and Assessment Authority is responsible for the setting of performance standards, the guiding of assessment design, and the moderating of assessment results.

Assessment Instruments for evaluating ApL subjects

Assessments of ApL subjects are designed, administered and judged by individual CPs. Students are not required to sit for public examinations. To enhance the comparability of standards across

ApL subjects, assessment principles, guidelines and instruments are required to have in place in both the pre-assessment and post-assessment stages. These serve to facilitate the setting and alignment of standards as well as the assessment decisions in the moderation process.

Assessment instruments are then designed based on the following guidelines:

- To align the formats and standards of Performance Descriptors for the Level 'Attained' among the 30 ApL subjects.
- To audit the coverage of Performance Descriptors against the Five Curriculum Pillars and the Intended Learning Outcomes (ILOs) pre-set in the ApL curriculum.
- To provide a common assessment framework for the development of ApL Assessment of different subjects.
- To check the appropriateness of the assessment tasks designed in which a wide range of abilities from reflecting the achievement of the ILOs can be assessed.
- To serve as a common tool to evaluate the cognitive level and skill coverage of the assessment tasks designed.
- To assist the assessors in making assessment decisions on the performance standards of students' works in the post-assessment moderation process.

Bloom's Taxonomy is used to classify different forms and levels of learning. Students' learning can be reflected through the outcomes of the assessment tasks. The instruments so designed are employed to evaluate these tasks with respect to different skills in the Cognitive Domain. The organisation and coverage of generic skills which are the core principles of the ApL curriculum framework can also be analysed in a similar manner. The results of the analysis may show how well an ApL subject meets the expected learning objectives.

Pre-Assessment Stage

(1) Setting of Standards and Performance Descriptors

The assessment results of ApL subjects will be recorded in the HKDSE diploma. Students' performance will be reported in two levels: 'Attained' and 'Attained with Distinction'.

A set of written descriptors has been developed for the 'Attained' level that describes what a typical student of a particular ApL subject performing at this level is able to do. The principle behind these descriptors is that they describe what typical candidates can do, not what they cannot do. These descriptors will necessarily represent 'on-average' statements and may not apply precisely to individual, whose performance within an ApL subject may vary. Samples of students' work at the 'Attained' level may be used to illustrate the standards expected of them. These samples, when used together with the written descriptors, explicitly illustrate the standards required at this level. In order to align the standards among subjects, performance descriptors of 'Attained' for each ApL subject have to be drafted covering the following seven dimensions:

- Knowledge and Understanding
- Application of Knowledge
- Generic Skills
- Communication Skills
- Subject-specific Performance related to the context
- Values and Attitudes towards the related industry
- Self-understanding for further studies and career development

The development of performance descriptors should start with the learning outcomes of the subject. It has to ensure that all learning outcomes need to be addressed by the seven dimensions

of performance descriptors developed. When drafting the descriptors, the use of descriptive words at the level 'Attained' have to be aligned for all 30 ApL subjects so that the students' performances among subjects can be compared. As a health check, a mapping table showing the relationship between learning outcomes and performance descriptors is expected. The importance of individual Learning Outcomes can also be revealed by inserting the degree of relevancy into the mapping table. Mapping of performance descriptors with the five curriculum pillars of Applied Learning further ensures comprehensive descriptions of students' performance in various foci of the curriculum.

Regarding the award of 'Attained with Distinction', a comparability analysis with the HKDSE core subjects (Chinese Language, English Language, Mathematics, and Liberal Studies) will be conducted. Those candidates awarded 'Attained with Distinction' are deemed to have performed at a level comparable to Level 3 or above under the standards-referenced reporting system of the HKDSE. The generic descriptors for the award of 'Attained with Distinction' have to be drafted comparable with the generic descriptors at Level 3 or above used by the core subjects. This methodology is to ensure that the ApL subjects are graded comparable with the core subjects in terms of the academic rigour.

(2) Assessment Framework of ApL Subjects and its Alignment in Design

Alignment of curriculum, learning and assessment plays a key role for the implementation of a coherent course. The Intended Learning Outcomes (ILOs) of the curriculum lead the assessment design as well as the relevant learning activities required. The ILOs of the curriculum, defining the outcomes to be achieved, encompassed the assessment dimensions to be assessed. Based on the identified assessment dimensions, assessment tasks with clear assessment criteria can be developed. The addition of differentiated descriptors to each assessment criteria further enhances the development of relevant rubrics for fair and valid judgment of students' performance. It also provides clear statements to students about their performance standards expected at various levels. Since the development of assessment tasks, criteria and rubrics are started with the ILOs, coherent assessment is expected. This backward design enforces the constructive alignment among ILOs, assessment criteria, assessment tasks, assessment methods and also the learning strategies required as shown in Figure 1.

To ensure a coherent assessment design from the Course Providers of ApL subjects, assessment literacy training, focusing on the basic concepts of Outcome-based Assessment (OBA), has been conducted. The key concept regarding the constructive alignment among curriculum, assessment and learning were emphasized. Assessment instruments for evaluating their assessment design, including mapping of assessment tasks with ILOs and assessment task specifications are introduced.

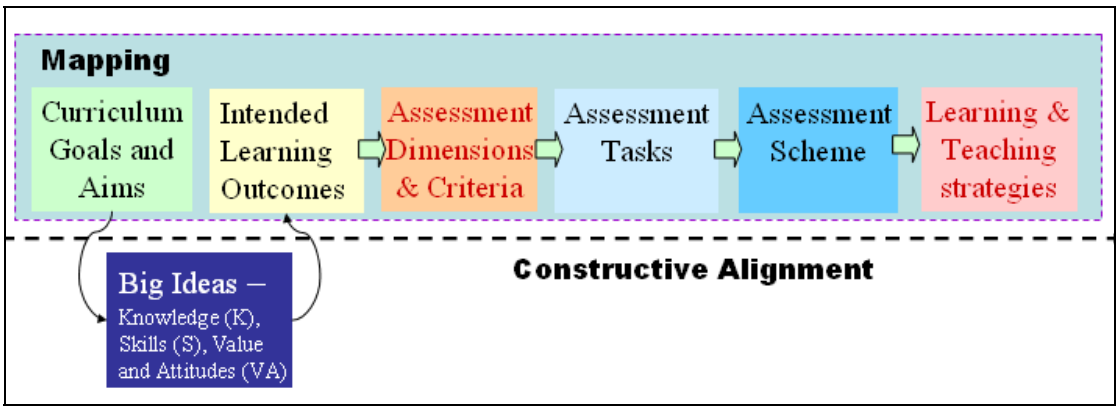

Figure 1: Constructive alignment among curriculum, learning and assessment

In the development of assessment scheme, CPs are required to state explicitly the dimensions to

be assessed for each assessment task. The assessment criteria under each dimension are further elaborated in the assessment task specification documents. The provision of assessment methods in the forms of rubrics, checklists and marking criteria further strengthens the articulation of judgement with those performances that are intended to be measured. As a health check for constructive alignment, assessment instrument such as mapping table of assessment tasks with ILOs is used. This mapping tool provides a holistic view on the relationship of assessment design and ILOs of the subject, looking for comprehensive coverage as well as the relative importance of ILOs revealed by frequency and weighting of the assessment tasks concerned.

Assessment reaches a high level of quality when it yields reliable, valid and useful information about students' performance (Carey, 2001). To ensure the validity and reliability, the assessment of ApL subjects should be designed in such a way that a wide range of abilities from reflecting the achievement of intended learning outcomes can be measured. To sustain the interpretation of attainment, the assessment design should consist of a representative set of tasks which can measure a wide spectrum of knowledge, skills and attributes, thereby leading to the requirement of multiple and varied assessments extending throughout the course. As ApL subjects are more performance-based and assessments usually take place over an extended period of time, a greater emphasis on formative assessment is expected. The formative assessment enhances learning through the provision of constructive feedback to the students while the summative assessment serves to conclude learning and teaching. Besides the coherence and the variety of tasks, the factors of authenticity, balance of theory and application, development of both practical and cognitive skills should also be duly considered in the design of ApL assessment.

ApL subjects are generally associated with a particular professional or vocational field. Apart from developing students' generic skills, assessments are primarily designed to meet both specific trade requirements and the expected learning outcomes of the subjects. Understandably, the assessment objectives, assessment methods and assessment criteria vary considerably across ApL subjects. Furthermore, students taking ApL subjects are not required to sit for the public examinations. Assessments are undertaken by individual CPs and administered by ApL subject tutors. Both factors impose difficulties in comparing the standards of works among different subjects. This creates the need for provision of a common framework and guiding principles in guiding the development of ApL assessment.

In response to the diversified nature of ApL subjects and the desire for high quality assessment, a common assessment framework in the form of assessment scheme has to be developed for each subject. This scheme spells out the details of what and how assessments should be carried out. In order to have better alignment among ApL subjects, the assessment framework is devised to include 8 to 10 assessment tasks in the scheme. When designing an assessment task, each of the following dimensions has to be considered:

- Assessment mode
- Assessment method
- Assessment criteria
- Assessment task weighting
- Weighting of individual and group assessment
- Weighting of in-class and outside-class assessment
- Weighting of written and practical assessment

The assessment scheme so designed for a particular ApL subject is then a main document for making comparison with others. In principle, well balanced in assessment modes and diversified assessment methods are expected to be found in the assessment scheme. These made assessment of ApL subjects reliable and valid. The 8 to 10 assessment tasks listed in the scheme form the major backbone of an ApL assessment. Based on the structure of this backbone, all ApL subjects can be made comparable.

In Applied Learning, it is intended that the assessment should stretch students' potential talents with opportunities to develop their knowledge and varied skills in different contexts. It is understandable that cognitive levels and generic skills coverage of assessment tasks are another important issue that needed to be considered. To address this issue, relevant assessment instruments have been developed. These instruments propose an Assessment Quality Index (AQI) used to assess the quality of assessment design of ApL subjects under the HKDSE system. The AQI is an index that measures the two main areas of assessment concern: cognitive requirements and generic skills coverage of the assessment tasks. Views of assessors from four essential fields, including curriculum developers, assessment experts, frontline tutors and subject experts, are collected to ensure comprehensive judgement from various perspectives. For each ApL subject, assessors are required to rate the assessment tasks in terms of relative opportunities for the development of various generic skills and the provision of various levels of cognitive challenges. Course Providers of ApL subjects are also requested to make use of these instruments for self-evaluation and seek for any improvement before the submission of the final Assessment scheme.

The development of AQI is a means that is used to maximize control on the quality of the assessment design and to quantify assessment decision in a more objective way. The availability of such index facilitates the comparison or benchmarking of the ApL subjects, regardless of their diversified nature. Deriving this index from the assessment instruments becomes an important activity in the pre-assessment stage for rating various designs of the ApL subjects.

Cognitive Level

In order to have the skills in cognitive domain assessed, cognitive level analysis tool will be used. All the 8 to 10 assessment tasks shown in the scheme will be analysed along the 6 levels of taxonomy: Knowledge, Comprehension, Application, Analysis, Synthesis and Evaluation. For each level of taxonomy, a degree of relevancy (0 to 3) is inserted so as to indicate the cognitive level of an assessment task. The aggregated score for each cognitive level forms an Assessment Quality Index of the ApL subject. Table 1 shows a matrix of the AQIs assessed by different stakeholders at different cognitive levels.

Cognitive Level Analysis of ApL Assessment Tasks - Subject X						
Assessment Quality Index - AQI_{cog}						
Assessors	L1 Knowledge	L2 Comprehension	L3 Application	L4 Analysis	L5 Synthesis	L6 Evaluation
CP	15	9	18	12	11	18
Assessment Expert (1)	19	8	11	12	12	7
Assessment Expert (2)	17	7	12	13	14	8
Curriculum Expert (1)	18	7	12	14	12	8
Curriculum Expert (2)	16	9	10	11	9	9
School Teacher (1)	17	6	12	12	11	5
School Teacher (2)	16	8	18	12	11	18
External Examiner (1)	18	9	12	12	9	9
External Examiner (2)	18	6	12	11	11	10
Mean	17.38	7.50	12.38	12.13	11.13	9.25
SD	1.06	1.20	2.39	0.99	1.64	3.85

Table 1: Assessment Quality Index for Cognitive Level Analysis

A radar chart (Figure 2) is further plotted with the AQIs stretched along the six cognitive levels so as to reveal the distribution of cognitive requirements of the assessment tasks in the scheme. The visual pattern of the chart provides a preliminary analysis of the assessment design in the cognitive domain.

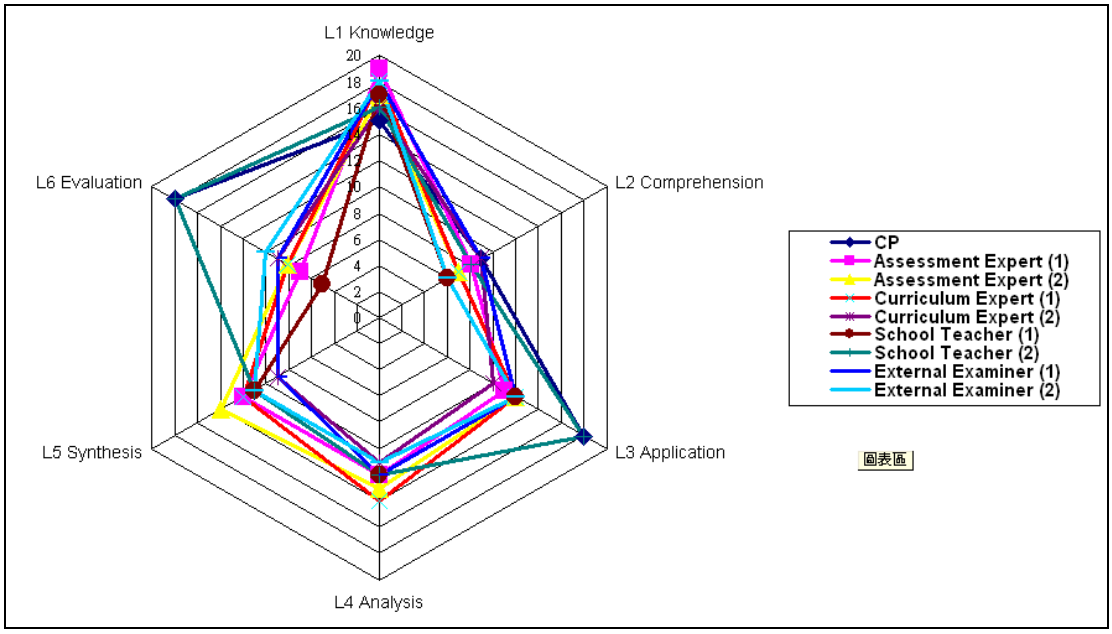

Figure 2: Radar Diagram of Cognitive Level Analysis

In general, the radar chart serves the following purposes:

- As a reflection tool for the Course Providers to further review their assessment designs.
- As an tool for the assessors to evaluate the cognitive level of the assessment tasks in the ApL subjects.
- As a common tool to compare the cognitive level of ApL subjects with similar nature (within the same area of studies).
- As a common tool to compare the cognitive level of ApL subjects with different nature (across area of studies).

The standard deviation (SD) of AQIs for each cognitive level may reveal the coherence of views from assessors of different fields and may evoke further points for discussion. It is also possible to individually compare the means across different subjects and identify any specific patterns obtained. A calculated overall mean can finally serve as an important index for making comparison among different ApL subjects.

In general, if the patterns of charts from different assessors are similar and the SD values of AQIs for various levels are small, it means that the assessors share common views on the cognitive levels of the assessment tasks. Instead, if the patterns of charts show significant discrepancies, actions may be needed to analyse the assessment design further. The use of these assessment instruments, AQIs and radar charts, provides a common platform to compare the cognitive levels of the assessment tasks of different ApL subjects. Consequently, the alignment of standards of ApL subjects in different clusters or areas can be achieved.

Generic skills

In addition to the cognitive requirements of the Assessment design, the generic skills coverage is another major concern to be addressed. Assessment instruments for checking skills coverage need to be completed by assessors of different fields. A similar Assessment Quality Index (AQI) is devised to mark the relevancy of the tasks related to different generic skills. Table 2 shows a matrix of the AQIs assessed along the ten dimensions of the generic skills. High AQIs are observed under the dimension of 'creative thinking' showing that this ApL subject is likely to be a subject in the Area of Creative Studies. The results of the matrix can also be plotted on

similar radar charts as shown in Figure 3. Again, a visual pattern can easily be seen. In general, similar chart patterns are expected for ApL subjects of similar nature.

Generic Skills Analysis of ApL Assessment Tasks - Subject X										
Assessment Quality Index for Generic Skills- AQI_{skills}										
Assessors	Communication skills	Numeric skills	Information skills	Problem-solving & decision making skills	Analytical skills	Creative thinking	Understanding interdependency & relationships	Self-reflection & self management skills	Interpersonal skills	Collaborative and Team Building Skills
CP	15	2	19	16	8	18	9	15	8	8
Assessment Expert (1)	14	2	16	8	15	17	9	10	5	7
Assessment Expert (2)	13	2	15	11	14	17	8	11	6	6
Curriculum Expert (1)	11	1	17	9	12	18	9	12	5	6
Curriculum Expert (2)	11	1	12	9	12	18	8	12	5	6
School Teacher (1)	11	2	12	9	12	17	8	12	6	6
School Teacher (2)	12	3	12	8	12	17	7	10	5	6
External Examiner (1)	13	2	14	10	11	17	8	12	7	7
External Examiner (2)	10	2	15	11	13	17	7	12	7	7
Mean of AQI_{skills}	11.88	1.88	14.13	9.38	12.63	17.25	8.00	11.38	5.75	6.38
SD of AQI_{skills}	1.36	0.64	1.96	1.19	1.30	0.46	0.76	0.92	0.89	0.52

Table 2: Assessment Quality Index for Generic Skills Analysis

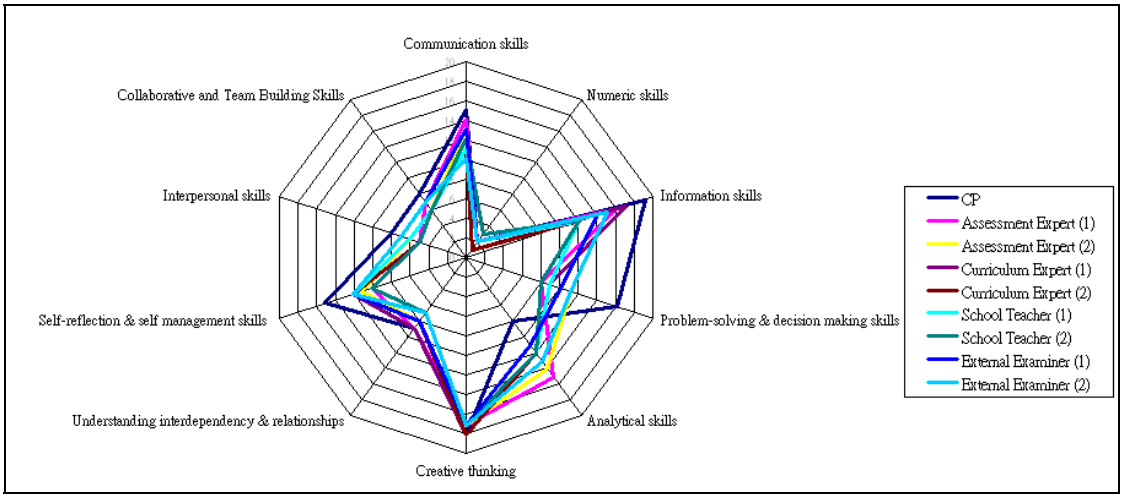

Figure 3: Radar Diagram of Generic Skills Analysis

Post Assessment Stage

For ApL subjects, assessments are designed, administered and judged by individual CPs. Students are not required to sit for public examinations as they do in Category A and Category C subjects under the HKDSE system. To enhance the comparability of performance standards across ApL subjects, the assessment results are subject to moderation by expert panels appointed by the HKEAA, and ultimately to approval by the HKEAA Public Examinations Board.

The moderation process involves the reviewing of subject assessment results and selected samples of students' work by a moderation panel. In this process, assessment criteria, achievement standards and adherence to the assessment framework will be taken into consideration. To facilitate the moderation judgement, a common assessment tool has been developed for the assessors to rate the performance standards of the students' work. This analysis tool serves as an ordinal measurement which complements the holistic expert judgement on the overall standards of work against the pre-determined standards as specified in the performance descriptors. It also establishes the common platform for comparing the standards of work across different subjects in the Area Moderation process.

Conclusion

Applied Learning subjects have a diversified nature. They have to meet societal needs and the community's expectation. Every year, some subjects may go obsolete while some new subjects may emerge. A review procedure is needed for the development of new ApL subjects so as to ensure their standards and the smooth implementation of the assessment strategy. The instruments or tools suggested in this paper may help to monitor the growth of ApL subjects in a systematic way.

References

Education and Manpower Bureau (August 2006) *Action for the Future --- Career-oriented Studies and the New Senior Secondary Academic Structure for Special schools,* HKSAR, PRC

Education and Manpower Bureau (May 2005) *The New Academic Structure for Senior Secondary Education and Higher Education --- Action Plan for Investigating in the Future of Hong Kong,* HKSAR, PRC

Curriculum Development Council (June 2001) *The Way Forward in Curriculum Development --- Learning to Learn,* HKSAR, PRC

Education Commission (September 2000) *Learning for Life Learning through Life --- Reform Proposals for the Education System in Hong Kong,* HKSAR, PRC

The Curriculum Development Council and the Hong Kong Examinations and Assessment Authority (2009) *Applied Learning Curriculum and Assessment guide (Senior Secondary Level)*

FU, T.W., & NG, K.M. (2008, October) *Development of Assessment Framework of Applied Learning Subjects in Hong Kong.* Paper presented at the 2008 IAEA Conference, Cambridge, United Kingdom

Biggs J. (1999) *Teaching for Quality Learning at University,* Society for Research into Higher Education/ Open University Press, Buckingham

Biggs, J (1996) *Enhancing teaching through constructive alignment*. Higher Education, 32,347-364.

Bloom B.S. (1964) *Taxonomy of Educational Objectives,* David Mckay

Michael Friendly (1991). *"Statistical Graphics for Multivariate Data"*. Paper presented at the SAS SUGI 16 Conference, Apr, 1991

Wiggins, G & McTighe, J. (2005). Understanding by design (2nd edition). Alexandria, VA: Association for Supervision and Curriculum Development

Wiggins, G (1997, September). *Work standards: Why we need standards for instructional and assessment design*. NASSP Bulletin, 81(590), 56-64

Anderson, L. W., & Krathwohl, D.R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* New York: Longman