

Analysis of the AIC Associate Constructor Practice Exam for Cognitive Domain and ACCE Student Learning Outcomes (SLO) Inclusion

Tom M. Leathem, Ph.D., Eric M. Wetzel, Ph.D., Megan Good, Ph.D., Katie Boyd, Ph.D., and Richard Burt, Ph.D.
Auburn University
Auburn, AL

Identifying direct assessments appropriate to the requirements for the American Council of Construction Education (ACCE) has become a topic of increased interest among construction educators. The Associate Constructor exam is one such direct assessment under consideration. Although the AC exam has been validated as a mechanism for the certification of constructors, use of the exam for accreditation purposes requires appropriate validation achieved through extensive evaluation of the content, construct, and criterion parameters of the exam. As a first step, the intent of this research was to identify the topical and cognitive scope of the AC practice exam in relation to the ACCE SLO requirements. Analysis of the practice exam questions showed that the test generally covers the range of subject matter and topical content associated with the discipline, but falls short of the appropriate cognitive levels in many areas per ACCE requirements. This paper provides a background and basis for the study, description of the methodology, results, and suggestions for further research of the full AC Level I exam to determine its validity as measurement tool for ACCE accreditation purposes.

Key Words: ACCE, AC Exam, Assessment, Bloom's Taxonomy, Student Learning Outcomes

Introduction and Background

The American Council for Construction Education (ACCE) requires all construction programs to assess programmatic student learning outcomes (America Council for Construction Education, 2014). The purpose of this assessment is to continuously improve the program; indeed, each program must also submit a quality improvement plan to ACCE based on assessment findings (ACCE, 2014). The standard dictates that at least one of these must be a direct assessment. ACCE defines a direct assessment as "…a student product or performance that can be evaluated" (ACCE, 2016). This move has caused many construction programs to consider different types of assessments to meet the accreditation requirement.

The Associate Constructor (AC) exam, developed by the American Institute of Constructors (AIC) Constructor Certification Commission (CCC) is one such direct assessment under consideration. Previous research has indicated that many construction programs are already utilizing this exam to measure student competencies (Olsen, Taylor, & Holk, 2011; Sylvester, 2011). AIC reports that about 50 construction management programs currently use the exam and many of those require it as part of their program (American Institute of Constructors Constructor Certification Commission, 2017b). Certification through the exam is identified to provide an individual "formal recognition of the education and experience that defines you as a professional…in the construction industry" (AICCCC, 2017a). The level one handbook for candidates (2017a) suggests some of the specific benefits of the AC certification are: certification of construction management skills and knowledge, analysis of individual strengths and weaknesses, and an independent assessment of skills and knowledge.

The AC exam is the first of two levels of certification offered by AIC. It consists of 300 multiple-choice questions administered over one day in two four-hour blocks. Skills and knowledge are measured and reported within the subject areas of: 1. Communication Skills, 2. Engineering Concepts, 3. Management Concepts, 4. Materials Methods Project Modeling and Visualization, 5. Bidding and Estimating, 6. Budgeting and Cost Control, 7. Planning and Scheduling, 8. Construction Safety, 9. Construction Geomatics, and 10. Project Administration.

As a means to help prepare potential test-takers for the exam, AIC provides an official study guide. The guide's topical information is organized into chapters by the different subject areas with review questions at the end of each chapter. In addition, there is a 100 question practice exam included at the end of the study guide. According to the instructions for the practice test, the questions are "similar to the ones you will find on the Level 1 Associate Constructor Certification Exam. The number of questions in each section of the practice test is proportional to the weight of each subject category." (AICCCC, 2017b, p. 153).

Research Aim and Objectives

Given the need to identify assessment tools to meet recent ACCE accreditation requirements and the current use of the AC exam by CM programs, utilization of the exam for accreditation purposes is a logical consideration. However, the exam was not originally developed for this purpose. As such, in-depth analysis of the exam would be necessary to identify its appropriateness for use in demonstrating student achievement of ACCE SLOs. This type of evaluation requires extensive validation study, which is beyond the scope and intent of the research in this paper. As a first step, the intent of this research was to identify the topical and cognitive scope of the AC exam in relation to the ACCE Student Learning Outcomes. The aim was to analyze the 100 questions of the practice exam included in the Associate Constructor Exam Official Study Guide (AICCCC, 2017b). Analysis of the questions was guided by the following objectives:

Objective 1: Identify the knowledge subject matter and/or topical content contained in each question on the practice exam.
Objective 2: Identify and categorize the cognitive domain embedded in each exam question relevant to the six Bloom's Revised Taxonomy domains of Create, Evaluate, Analyze, Apply, Understand, Remember.
Objective 3: Categorize the identified subject matter/topical content into the relevant ACCE Student Learning Outcomes.

Literature Review

The current Standards for Educational and Psychological Testing state that "validity refers to the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests" (AERA, APA, NCME, 2014, p. 11). That is, a test cannot be validated; rather, validity evidence can be gathered to support *interpretations* of test scores for particular purposes. Indeed, the Standards go on to say that "Evidence of the validity of a given interpretation of test scores for a specific use is a necessary condition for the *justifiable use of the test*" (AERA, APA, NCME, 2014, p.11, emphasis added).

When the Associate Constructor Certification (Level 1) Exam was created, a great deal of attention was given to collecting validity evidence with the purpose to "…ensure that the examinations reflect the knowledge and skills necessary for competent performance in managing the construction process" (American Institute of Constructors, 2013). The test creators used the Standards for Educational and Psychological Testing (AERA, APA, NCME, 1985) as a basis for test construction. Of note, the 1985 standards also viewed validity as a process focused on providing validity evidence for specific *inferences* as opposed to collecting validity evidence for a test.

The test developers of the Associate Constructor Certification (Level 1) Exam, used an adapted version of Haladyna's Mental Behavior Classification System (1997) to determine the cognitive levels the test would measure. This model asked subject matter experts to create items that aligned with the following levels of learning: Recall, Understanding, Problem Solving, Evaluating, and Predicting. Thus, the test was carefully designed for its intended purpose and psychometric evidence was gathered to support the certification of entry-level Building Construction

Managers. However, it does not appear that validity evidence was collected to support test inferences for programmatic learning improvement purposes. If the AC exam is to be used for this purpose, then new validity evidence is needed to support these inferences.

When reviewing assessments for validity there are generally three areas to consider - content, criterion, and construct. Each area plays a part in determining the level to which we can rely on the inferences made from the assessment data. Content-related evidence of validity is arguably the most critical part of assessment validity. This area refers to the alignment of content included in the assessment relative to the identified aim with which it intends to measure (Popham, 2014; Wiggins, 1998; Wiggins & McTighe, 2005). Criterion-referenced evidence of validity relates to assessments that are used to generate inferences about a student's ability to succeed at subsequent tasks (Popham, 2014). Construct-related evidence of validity references an assessments ability to measure what it purports to measure (Russell & Airasian, 2012). The constructs of an assessment are the knowledge, skills, and/or performance to be exhibited by someone through execution of the assessment, such as the ability to interpret or recall.

Commonly in assessment, Bloom's Taxonomy is used to determine cognitive levels of learning (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1965). In a revised taxonomy by Anderson et al. (2001), a linear progression of cognitive levels of learning are presented: Remember, Understand, Apply, Analyze, Evaluate, and Create. The highest level of learning in the taxonomy is "Create" and the lowest level is "Remember". For the purpose of programmatic improvement, it is critical that test scores yield specific information that aligns with the desired levels of learning that can be used to make programmatic decisions.

Research Methods

Analysis of the AC Exam utilized a systematic coding of each of the 100 practice exam questions to identify the ACCE SLO and revised Bloom's cognitive domain. All researchers involved in the study hold PhD's, one with a concentration in assessment, and all with prior experience in assessment evaluation. As a second measure to ensure robustness of the study, the methodology and results were reviewed and validated by university assessment experts. The following research questions were constructed to guide the research for data collection and analysis.

Research Questions

1. For each question on the practice exam:
    a. Which ACCE SLO(s) is the question addressing?
    b. What is the cognitive domain that the question reaches, based on Anderson, et al revised Bloom's Taxonomy?
2. What percentage of questions aligned within an ACCE SLO meet or exceed the ACCE minimum identified cognitive level?

Research Approach

The research utilized a systematic, phased approach in order to obtain, analyze, and validate the research findings. The three phased approach allowed for a validation checkpoint at the end of each phase. Phase I was a pilot study in order to validate the intended approach of the research. Phase II executed the data collection and initial analysis. Finally, Phase III finalized the analysis and utilized a third party assessment expert to validate the results. *Figure 1* presents the phased approach.

Phase I

Phase I of the research was used in order to test the coding process. During the pilot study, the three researchers randomly selected a group of 10 (out of the 100) AC practice exam questions to review. Individually, the researchers coded these 10 questions, identifying the cognitive domain and the ACCE SLO(s) that each question addressed. Each researcher placed their codes into a developed data collection matrix (Figure 2*).* Each ACCE SLO was included in the data collection matrix as the "knowledge domain" without its respective cognitive levels identified.

This was done purposefully to reduce the potential to bias the results on the identified cognitive level for each question.

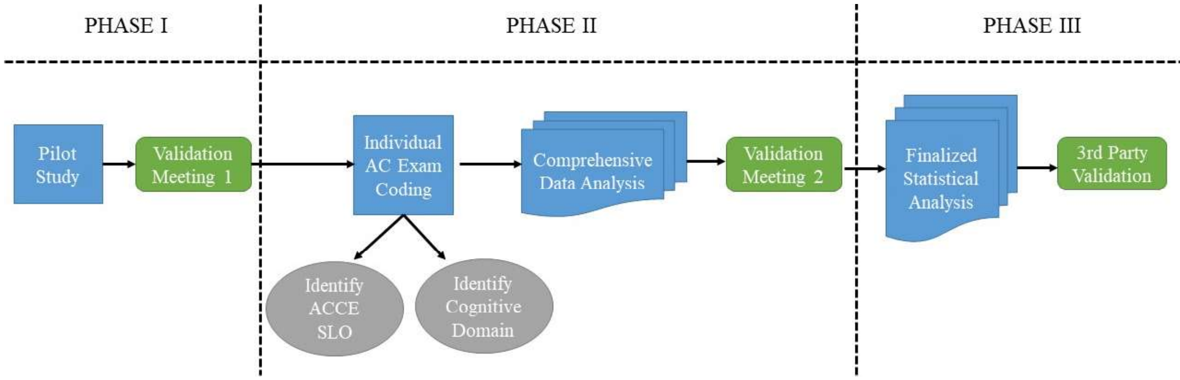

Figure 1: Phased Research Approach

The individual coding was combined into a master matrix and analyzed by one of the researchers to identify initial consensus and discrepancies among the three. Consensus was considered to be achieved when all three researchers agreed on an item. In order to eliminate discrepancies between researchers, a validation meeting was held to discuss where codes varied for the 10 pilot study questions. Consensus among the 10 questions was reached, validating the procedure for data collection.

	Bloom's Cognitive Domain and Associated Action Verbs																		
	Remember (R)		Understand (U)							Apply (AP)		Analyze (AN)			Evaluate (E)		Create (C)		
ACCE SLO Knowledge Domain	Recognizing	Recalling	Interpret	Exemplifying / Illustrating	Classifying	Summarizing	Inferring	Comparing	Explaining	Executing	Implementing / Using	Differentiating	Organizing	Attributing	Checking	Critiquing	Generating	Planning	Producing
1. Written communications appropriate to the construction discipline.		4																	
2. Oral presentations appropriate to the construction discipline.	81																		
3. Construction project safety plan.																			
4. Construction project cost estimates.		31								57									
18. Basic principles of sustainable construction.																			
19. Basic principles of structural behavior.																			
20. Basic principles of mechanical, electrical, and piping systems.			40																

Figure 2: Data Collection Matrix (Partial)

Phase II

Phase II utilized the same procedure as the pilot study, but analyzed the entire AC practice exam. Each researcher individually evaluated all 100 questions, identifying the ACCE SLO(s) and cognitive domain. The codes were compiled into a master matrix and analyzed by the same researcher as in the pilot. Then, a validation meeting was held by the research team to discuss, and come to consensus on, discrepancies between their individual results.

Phase III

Upon consensus for all 100 questions, the data was analyzed for statistical purposes. This step grouped the questions by SLO(s) and quantified the percentage of questions by cognitive domain. In order to validate the procedure and results, a third party from the university office of academic assessment was brought in to review the analysis consistency and reliability.

Results

Phase I

The pilot study included randomly selecting 10 of the 100 available AC practice exam questions provided in the AC exam study guide (AICCCC, 2017b). Each question was reviewed to identify key words indicating the cognitive domain and knowledge areas required (Figure 3).

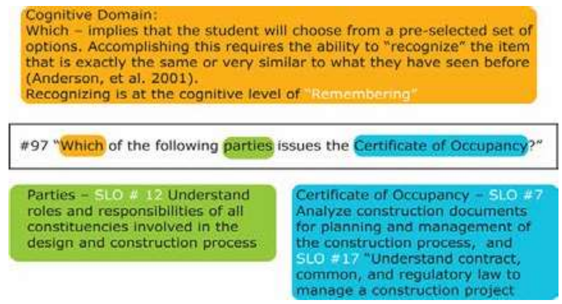

Figure 3: Example question analysis approach to identify cognitive and knowledge requirements.

Initial results of each researcher's coding identified a 70% and 90% consensus rate for the knowledge domain and cognitive domain, respectively (Table 1). After meeting to review the combined results and discuss the three questions with identified discrepancies, consensus was able to be reached on all 10 questions. Question 25 was identified to be more of a general curriculum question and a great deal of discussion was had about how to address questions of this nature. It was agreed by the team to try and identify the most closely relatable SLO for each question but acknowledge that successfully answering the question requires knowledge mostly outside of the construction discipline. Based on the final results, the methodology was validated and Phase II commenced.

Table 1

Phase I Analysis Results

AC Exam Question	ACCE SLO #			Cognitive Domain (CD)			Result	
	R1	R2	R3	R1	R2	R3	SLO(s)	CD
4	7, 9	1, 7	7	R	R	R	7	R
9	6, 9	9	9	R	R	R	9	R
20	7, 12, 14, 17	17	17	R	R	R	17	R
25*	10	7	1, 7	R	R	R	7[1,2]	R
31	4, 8	8	8	R	R	R	8	R
40	7, 20	7	7	U	U	U	7	U
57**	4, 14, 16	7, 14, 16	5	AP	AN	AP	4,5,14,16[1,3]	AP[1]
68	5, 12, 15	12	12	R	R	R	12	R
81	3	3, 13	3	R	R	R	3	R
92*	17	16	17	R	R	R	17[1]	R

Note: *= discrepancies on SLO needing further review. **= discrepancies on SLO and cognitive domain needing further review. [1] consensus result after review meeting. [2] Most closely relatable SLO but question is mostly non-construction related knowledge. [3] question requires multiple areas of knowledge because it is a higher level cognitive ability.

Phase II

All questions included in the AC practice exam were reviewed and coded by each researcher independently to identify the related ACCE SLOs and cognitive domain for each question. Care was taken to ensure each researcher's analysis results were isolated from the other researchers to reduce potential bias. Percent agreement among the researchers was calculated as a measure of inter-rater reliability, which is to say the percent of questions by which the three independent raters categorized the cognitive level of the exam questions the same. It can be quite difficult to obtain a perfect match or alignment when evaluating the percent agreement with only two independent raters, let alone three or more. Additionally, the more categories available to the raters, the more difficult it becomes to obtain high percent agreement (e.g., the raters coded each question of the practice exam as one of six cognitive levels: “Remember”, “Understand”, “Apply”, “Analyze”, “Evaluate”, or “Create”). If there was absence of agreement by the raters, then by chance alone the percent agreement would be less than 3% for each question (when you have three raters and six options). Across the entire practice exam, the percent perfect agreement was 58.2%. It should be noted that of the 100 questions, one of the questions were duplicates and therefore the reliability used N=99. Of note, the raters all independently agreed there were not questions on the test that should be categorized as “Evaluate” or “Create”, the two highest cognitive levels of learning. As a result the range between levels of ratings across each question was never larger than four (“Remember”-“Understand”-“Apply”-“Analyze”).

As previously noted, perfect agreement calculations employ a rigorous standard for reliability. Often when calculating inter-rater reliability, the *adjacent* agreement is also calculated. Here, a tolerance of 1 was applied (i.e., raters were within 1 point of agreement of each other). The adjacent agreement in this study was 91.8%. That is, the raters were fairly consistent in rating the cognitive level of each exam question and over 90% of the time, were within just one rating level from each other. Finally, it is important to note that disagreement between raters was contained to just 29 of the questions. Raters displayed perfect agreement across 69 of the 99 questions on the exam, in regards to cognitive level.

Results of the consensus analysis identified a total of 56 questions where consensus was reached in the initial coding. Details are illustrated in figure 4. Black numbers identify questions with consensus, red indicates the number of questions lacking consensus, and light gray bars identify the range of cognitive ability identified for all questions regardless of consensus status. Out of the 44 questions needing further review, 19 showed discrepancies on the knowledge and cognitive domains, 15 showed differences in only the knowledge domain, and the remaining 10 differences were in the cognitive domain. Similar to the pilot phase, the majority of questions (84%) containing discrepancy among both domains were scored at the “apply” or “analyze” cognitive domains. The propensity for discrepancy at this level is reasonably greater given the requirement for more areas of knowledge to solve problems at these higher levels of ability. A follow-up meeting to discuss the merits of each researchers scoring for the 44 questions resulted in consensus on both domains for the remaining questions.

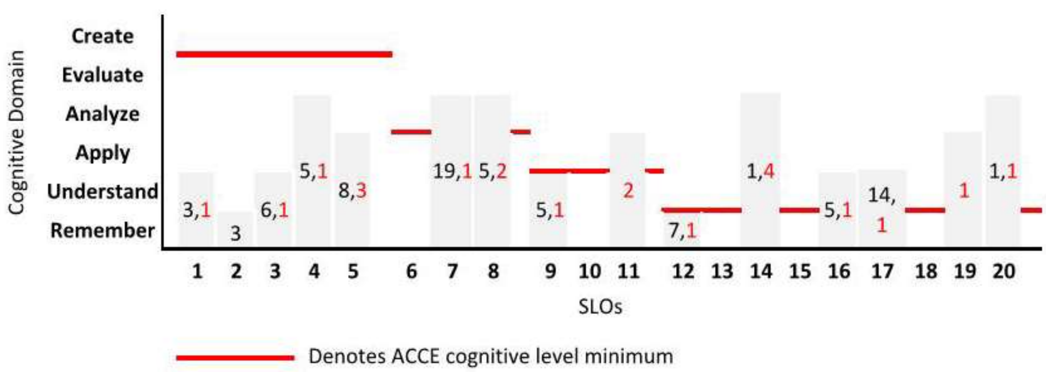

Figure 4: Initial question (N=99) analysis results identified by numbers of questions identified for each outcome and cognitive domain.

Phase III

After reaching consensus on all 100 questions, the codes were aggregated to identify the total number of questions that relate to each ACCE SLO and their associated cognitive domain (Figure 5). Results indicate that 16 of the 20 SLOs are being addressed to some extent by the exam questions. Out of the 100 questions, 25 met or exceeded the minimum cognitive level identified in each SLO. Three quarters (N=20) of those were associated with SLOs 12-20. The remaining five questions were distributed among SLOs 6-8 (N=3) and 9-11 (N=2). All of the questions (N=52) that were aligned to SLOs one through five – ranked at the “Create” cognitive level – fell short of reaching the minimum level ACCE requirement. This is not surprising considering accurate assessment of cognitive abilities at this level are extremely difficult, if not impossible using item-response assessments (Leathem, 2017; Wiggins, 1998; Bloom, et al., 1956).

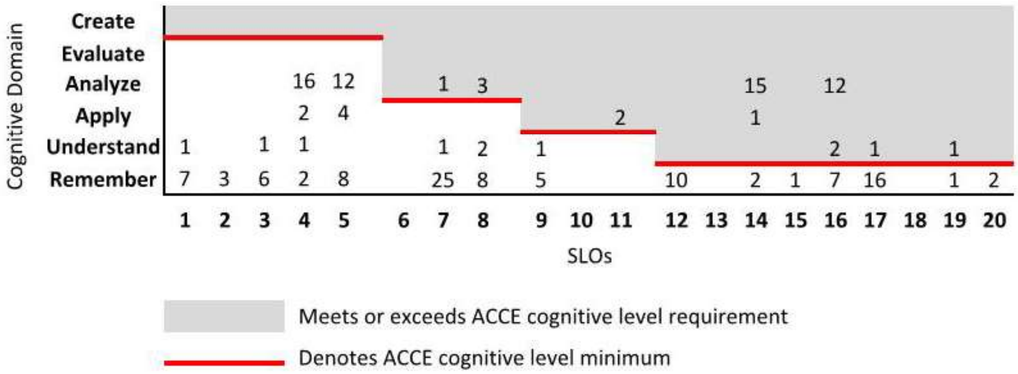

Figure 5: Number of exam questions categorized by outcome and cognitive requirement.

Discussion and Conclusion

Construction programs are seeking appropriate examples of direct assessment tools to evaluate their programmatic outcomes to meet ACCE accreditation requirements. The AC exam is one such instrument under consideration. Because the AC exam was not originally developed for this purpose, appropriate evaluations need to be conducted to establish its validity. As a common first step in validity evaluation (Wiggins, 1998), the aim of this research was to analyze the 100 questions in the AC practice exam to identify the scope of knowledge and cognitive inclusion in the exam based on Bloom's Revised Taxonomy.

Results of the study suggest the test covers 80% of the range of subject matter and topical content as identified by the ACCE SLOs. However, 35% (N=7) of the SLOs are included at or above the minimum cognitive levels per ACCE requirements. The majority of those fall within SLOs 12-20, at the "understand" cognitive level. This is understandable considering an item response type instrument like the AC exam is more typically suited for this level of evaluation. Further, of the (N=20) questions in this area, 18 of the questions are confined to SLOs 14 and 16, leaving minimal inclusion to the other SLOs. In three cases, only one question is included at or above the minimum cognitive level. This is not to suggest that the other SLOs are inadequately included. The depths of this particular study were too limited to make that type of inference. Rather, it suggests a potential opportunity to create a more balanced inclusion of questions that would be more evenly distributed across SLOs 12-20.

Analysis of the practice exam provides a meaningful first look at the full AC Level I exam considering AIC indicates the practice exam is representative of the breadth and depth of the full AC exam. However, the authors recognize that because this study was limited to the practice exam, generalizability of the results may not be fully representative of the AC exam. Analyses of the full exam using the methodology in this study may yield different results. Recognizing this potential limitation, the authors recommend that further research be conducted using actual questions used in the official AC exam to gain a more complete picture of the content and cognitive inclusion of the actual exam as it relates to the ACCE SLOs. As well, further research needs to be conducted to consider other elements of the exam to establish a complete conclusion about the validity of the AC exam as a direct assessment for accreditation purposes.

References

American Council for Construction Education. (2016). Document 103 OBS (Revision 08/2016): Standards and criteria for accreditation of postsecondary construction education degree programs. Retrieved from http://www.acce-hq.org/images/uploads/Doc_103_Updated_081116_final.pdf

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing. (2014). *Standards for educational and psychological testing.* Washington, DC: AERA.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing. (1985). *Standards for educational and psychological testing.* Washington, DC: AERA.

American Institute of Constructors (2013). *Qualitative and Quantitative Evidence Collected for Validation of the Constructor Certification Process.* (No. 17). American Institute of Constructors. Retrieved October 15, 2017, from https://c.ymcdn.com/sites/professionalconstructor.site-ym.com/resource/collection/A5D72363-A21B-423B-88F2-A820F54DACB9/Commission%20Document%20No.%2017%20-%20Qualitative%20and%20Q.pdf

American Institute of Constructors (2013). *Examination Item Writing Form.* (No. 18). American Institute of Constructors. Retrieved October 15, 2017, from https://professionalconstructor.site-ym.com/general/custom.asp?page=PoliciesandDocuments&DGPCrSrt=&DGPCrPg=2

American Institute of Constructors Construction Certification Commission. (2017a). Certification examination for level I associate constructor: Handbook for candidates. Retrieved October 6, 2017 from: http://c.ymcdn.com/sites/professionalconstructor.site-ym.com/resource/collection/7E04C2B6-4A57-4F62-B556-5A1EB71BB9C6/20170117_AICCCC_Level1_Handbook_2017.pdf

American Institute of Constructors Construction Certification Commission. (2017b). Associate constructor level 1 official study guide: Revised September 2017. Retrieved October 6, 2017 from: http://c.ymcdn.com/sites/professionalconstructor.site-ym.com/resource/collection/7E04C2B6-4A57-4F62-B556-5A1EB71BB9C6/AC_Exam_Study_Guide_Revised_September_2017.pdf

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 (Complete Edition), Longman, New York, NY.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives, handbook I: The cognitive domain* (Vol. 19, p. 56). New York: David McKay Co Inc.

Haladyna, T. M. (1997). *Writing Test Items To Evaluate Higher Order Thinking*. Allyn & Bacon, 160 Gould Street, Needham Heights, MA 02194-2310.

Leathem, T.M. (2017). *A curriculum review of a construction management program in consideration of accreditation by the American Council for Construction Education* (Doctoral Dissertation). Retrieved from Mississippi State University Dissertation Library.

Taylor, J., Olsen, D., and Holk, J. (2011). Do professional credential supported by ASC member schools focus on theose most likely to enhance a students' professional development?. *Associated Schools of Construction 47th International Proceedings.*

Popham, W.J. (2014). *Classroom assessment: What teachers need to know*. Upper Saddle River, NJ: Pearson

Sylvester, K. (2011). Using the AIC level I examination to assess student learning. *Associated Schools of Construction 47th International Proceedings.*

Russell, M.K. & Airasian, P.W. (2012). Classroom assessment: Concepts and applications *(7th Edition)*. New York, NY: McGraw-Hill

Wiggins, G.P. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco, CA: Jossey-Bass.

Wiggins, G.P., & McTighe, J., (2005). *Understanding by design (expanded 2nd ed).* Alexandira, VA: Association for Supervision and Curriculum Development.