Using Secondary Evaluation to Assess Equivalency of Learning across Various Curriculum Delivery Modalities

Kelly Whealan George  
*Embry-Riddle Aeronautical University, georged8@erau.edu*

Debra T. Bourdeau  
*Embry-Riddle Aeronautical University, taylo13f@erau.edu*

Constantine M. Koursaris  
*Embry-Riddle Aeronautical University - Worldwide, koursarc@erau.edu*

Follow this and additional works at: [https://commons.erau.edu/publication](https://commons.erau.edu/publication)

Part of the [Higher Education Commons](https://commons.erau.edu/publication)

Scholarly Commons Citation  

This Article is brought to you for free and open access by Scholarly Commons. It has been accepted for inclusion in Publications by an authorized administrator of Scholarly Commons. For more information, please contact commons@erau.edu, wolfe309@erau.edu.
Using Secondary Evaluation to Assess Equivalency of Learning across Various Curriculum Delivery Modalities

Kelly Whealan George, Ph.D
Department of Social Sciences and Economics

Debra Taylor Bourdeau, Ph.D
Department of English, Humanities & Communications
College of Arts and Sciences
Embry-Riddle Aeronautical University
600 South Clyde Morris Blvd.
Daytona Beach, FL 32114 USA

Constantine Michael Koursaris, Ph.D
Department of Decision Sciences
College of Business
Embry-Riddle Aeronautical University
600 South Clyde Morris Blvd.
Daytona Beach, FL 32114 USA

Abstract
This study investigates whether research-based student learning outcomes (SLOs) are equivalent across various curriculum modalities. The study was conducted to assess student artifacts collected from undergraduate courses containing a gradable research-based assignment activity. The researchers of this study either led or participated in the effort to re-evaluate these graded assignments in an unbiased, verifiable approach and score their findings using levels of attainment or competency on a scale from zero to four, with zero being at the lowest and four being at the highest levels of competency. The raw data gathered were then tabulated and analyzed using statistical significance of the results that aided in answering the research questions posed in this study. The researchers hope that the results of this study will be useful in the consistent delivery of equivalent courses across various learning modalities.

Keywords: Evaluation, Assessment, Learning, Program, Outcome, Curriculum, Higher Education

1. Introduction
As the delivery of education curricula shifts between synchronous and asynchronous learning styles, the quality of education, along with knowledge gained, is expected to remain transparent and equivalent regardless of curriculum delivery modality. According to Wedlock and Growe (2017) the field of education has yet to structure teaching and learning to match the coming digital generations. It is within every academic institution’s vast interest to strive and structure an equivalent transparent implementation of fulfilling student learning outcomes regardless the delivery modes of both traditional classroom lectures and online settings. Gardner (2017) reports that students need to develop skills and conduct research in order to become employable. The intent of this paper is to evaluate and assess previously submitted student research artifacts to determine, whether course learning outcomes, match the established university wide research student learning outcomes. Additionally, if student experiences are structured effectively and processed rigorously, they can add a great deal of value to students’ learning and to the educational strength of the university (Stirling& Kerr, 2015).
1.1 Scenario
The researchers of this study are full-time faculty at Embry-Riddle Aeronautical University, possessing extensive experience of collecting substantial outcomes-based data for performing course and program assessments to improve student learning outcomes and experiences while satisfying accreditation requirements. The university features three campuses; two residential campuses and a distributed “worldwide” campus.

The distributed-model campus includes on-campus, face-to-face (f2f) learning in centers around the world as well as online or hybrid courses. The various course modalities, when combined, resulted in over 85,000 student registrations and a headcount of almost 22,000 students for the 2015-2016 academic year. Campuses are located in both civilian (such as Atlanta or Orlando) and military (such as Joint Base Andrews or Minot Air Force Base) locations, and students are almost equally classified as military (10,878 in 2015-16) and civilian (11,070 in 2015-16). Students can choose from programs in three academic colleges, each with a mixture of full-time, and adjunct faculty, with a heavy reliance on adjunct faculty. Students may also choose a mixture of course delivery modalities in pursuit of completing their degree program.

Courses are offered in nine-week terms. Five “main” terms, including face-to-facelecture and hybrid courses, begin in January, March, May, August and October. Online courses begin each month. The wide distribution of locations, accelerated term schedule, frequency with which courses are taught, and reliance on adjunct faculty, have raised internal concerns about the consistency of the overall student learning experience and effective competency achievement of acquired knowledge, skills, and abilities. These concerns are also echoed at other academic institutions and regional accrediting bodies in the higher education industry. As a result, several steps have been taken to both create this consistency and to assess how well these efforts have fared as part of a continuous improvement effort in the quality of education provided to our students holistically.

1.2 Creating a Consistent Experience
The university has three distinct delivery modalities, each with a unique set of features that appeal to specific populations of students. There are, of course, traditional lecture courses where a faculty member teaches to a room of students synchronously. Additionally, students can learn synchronously by attending class from home using web conferencing technologies. This modality, called EagleVision home, allows students who may not have convenient access to a physical campus location to experience live instruction. A hybrid of these modalities, called EagleVision classroom, features a “host” campus (where the instructor is present before students) and “remote” campuses, where students in classrooms participate in class virtually using cameras and microphones. The EagleVision classroom courses allow several smaller sections to combine and create a larger, viable section that better meets minimum enrollment standards.

Either synchronous option can be “blended” with online instruction to provide alternatives for meeting required course contact time. Instructors often use peer reviews, discussion boards and other learning management system (LMS)-based activities as part of these blended courses. These courses require a more extensive use of the LMS, with requirements for students and instructors to interact each week.

There are also completely online, asynchronous courses. These courses are templated, and all sections of a course feature the same assignments developed by a faculty member, who is a subject-matter expert, in concert with an assigned instructional designer. Both work to ensure that the course meets learning (LO) and program outcomes (PO) by careful mapping of each assignment. In fact, the alignment of each element of the course to its module, learning and program outcome is so thorough and rigorously examined that elements of online courses are required in the other modalities.

These required components form what is called the Multi-Modality Template (MMT) and help to ensure consistency and aid formal assessment by creating a common set of assignments (and associated grade weights) that will appear in every section of the course regardless of modality. MMT assignments are selected to include a broad coverage of the course learning outcomes while preserving the academic freedom of instructors to deploy their own preferred materials in the course. Because MMT assignments are selected to ensure LOs are being consistently met across such a distributed university system, they are the most obvious choice for inclusion in yearly academic assessment plans.

Embedded rubrics are another attempt to ensure course section consistency across delivery modalities while providing the secondary benefit of facilitating data-gathering in the assessment process. Most MMT assignments will feature a rubric embedded in the Canvas LMS.
Because instructors are required to use Canvas to record all grades, assessment coordinators can easily harvest the information from these rubrics in their attempts to determine the levels of achievement of student of learning outcomes. The use of common assignments, coupled with specified rubrics, should ideally create a consistent course experience whether the class is taken via pure lecture, EagleVision home or completely online. A student in Orlando, therefore, should have a comparable section of English Composition as a student in California or even at the Singapore campus.

This ideal, however, is mitigated by the reality that rubrics might be applied inconsistently (or even ignored) and that assignment instructions themselves may be widely and variously interpreted by instructors. This research was the first attempt at evaluating if the use of common assignments coupled with specified rubrics did indeed result in consistent student achievement of learning outcomes in a number of courses.

2. Literature Review

While much of examination of the efficacy of rubrics has focused on their varying reliability, Reddy and Andrade (2010) expressed concern that the quality of the rubrics themselves may impact how they are interpreted by instructors and called for a more intense study of rubric quality. Additionally, perceptions about the purpose of rubrics themselves can affect how willingly they are embraced by instructors. For example, an instructor who believes that the rubric is a tool solely for grading, rather than seeing their potential use in formative assessment, is likely to trust his or her own objectivity and be less interested in using rubrics. The “limited conception,” noted Reddy and Andrade (2010), can be contrasted to a student-centered perception of rubrics as “serving the purposes of learning and achievement” by clearly showing students various “targets” for optimal performance (p. 437).

Reddy and Andrade’s (2010) overview of published literature on rubrics and their usage also shows some evidence of improved academic performance and reports that students are able to demonstrate higher-level thinking, specifically when rubrics are used in self and peer evaluations. Simply handing the student a rubric, however, does not ensure success. Instead, the instructor must carefully explain its parameters and even, ideally, engage students in the composition of rubrics. Therefore, it is vital to erode the differences in instructor and student perceptions of rubrics.

In their study of attempts to develop a method of effectively evaluating documentation skills of medical school graduates, Boulet et al. (2004) determined that rubric reliability improved with training of raters. Their research showed a drastic drop in rating variance as training time increased. Although rater differences could not be completely eliminated, and there was some impact on scores based on the assigned rater, this effect is mitigated by employing multiple raters on the same artifact.

Norming exercises were also cited as essential by Turbow and Evener (2016) who calibrated their rubric in a series of workshops focused on discussion of the rubric criteria, application of the rubric to sample student work and the attempt to achieve inter-rater reliability. Raters confirmed that they were more confident in their scoring ability after the careful discussion of the items on the rubric and even “felt a greater level of comfort in communicating with one another” (Turbow & Evener, 2016, p.211). Raters, after the workshops, were able to successfully apply their skills to independent assessment of student work. There were issues, however, when applying the same rubric to a different type of assignment, indicating a need to clarify a relationship between the rubric and each specific assignment to which it applied.

Despite the consensus that calibration exercises improve rater reliability, the feasibility of such endeavors can be compromised when it is necessary on a large scale. Cash, Hamre, Pianta and Myers (2012) demonstrated, fortunately, that such an extensive calibration program is, in fact, possible. In their attempt to train over 2000 raters, the authors acknowledged that reliability is heavily dependent on both the context and the raters involved. Varying backgrounds, experience and even attitudes can impact inter-rater reliability. Additionally, they understand that budgets can be a concern, with variables including implementation costs of any program as well as the frequency of continued calibration efforts. Finally, finding opportunities to scale the project can be difficult, as trainer/rater availability can be a problem.

Cash, Hamre, Pianta and Myers (2012) observed Head Start’s efforts at large-scale calibration and their ultimate success. Training sessions were held regionally, improving attendance concerns. Training sessions included video segments to ensure consistency and included time for dialog and question-and-answer sessions to address misconceptions and concerns. Raters were invited to practice their skills and discuss the results.
These concerns about faculty and student perception of rubrics, their usage and quality and the clear necessity for frequent and consistent calibration of instructors are all valid issues that need to be addressed when attempting to discern the equivalency of student learning. These apprehensions are compounded when student learning takes place over a variety of geographical locations, modality types and with both full-time and adjunct faculty members. By using an evaluation of student artifacts by secondary evaluators, professors that are not involved in the specific instruction of the course, over samples from an entire year and all modalities hinted at possible inconsistencies of the student achievements of learning outcomes across time and modalities.

3. Methodology

This research was initiated from a question arising from the assessment activities required in the university’s Quality Enhancement Plan (QEP) required for Southern Association of Colleges and Schools (SACS) accreditation. In order to support our mission and vision, Embry-Riddle Aeronautical University implemented its QEP, known as Ignite, in academic year 2012-2013. This initiative sought an active learning environment dedicated to systematic inquiry to solve problems or to advance knowledge. The goals of the initiative were to deliver curriculum where students will obtain the skills to investigate hypotheses, solve problems, and advance knowledge utilizing various disciplines. The research-supportive curriculum provided undergraduates with a learning experience rooted in the process of discovery through research and inquiry. Implementation focused on undergraduates, following a tiered plan of introduction, practicing, and mastery of research skills within course-based research.

Ignite developed six student learning outcomes (SLO) to encompass the basic principles of research in every discipline. Infusing the SLOs into the curriculum enabled students to graduate with a strong foundation in research principles. The Ignite SLOs are as follows:

1. Define and/or articulate a research problem
2. Design a course of action to solve a research problem using appropriate multidisciplinary principles.
3. Apply ethical principles in research
4. Conduct research independently and/or collaboratively
5. Reach decisions or conclusions based on the analysis and synthesis of evidence
6. Communicate research results.

At the beginning of the QEP, faculty identified course activities that aligned to the Ignite SLOs one through six using the University Ignite Rubric (Appendix 1) as a guide for selecting research-related activities. This mapping primarily allowed leadership to track the evolution and desired increase in identified research-related activities and student performance within the curriculum. In addition, this selection also allowed the comparison of student performance across modalities because the artifacts were evaluated in the course and secondarily during the special assessment cycle using the rubric and faculty not involved in the original grading of assignments. Therefore, the same student artifacts for the SACS assessment were used as the basis for this research. Assignments that were used for the SACS assessment project targeted frequently taught courses as well as the proficiency level of research expected from the assignment. The population of identified assignments was student artifacts used in this research were submitted during the academic year 2014-2015. From this population, the student artifacts were collected from one class section during every term the course was offered. If there was more than one modality offered in one term, student artifacts were collected from one course in each modality to ensure representation of all curriculum delivery modality. Table 1 indicates the course artifact, total number of submissions from each delivery modality and intended competency level.

<table>
<thead>
<tr>
<th>Course</th>
<th>Total Number of Artifacts</th>
<th>Online or Traditional</th>
<th>Competency Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Management 391</td>
<td>70</td>
<td>58</td>
<td>12</td>
</tr>
<tr>
<td>Safety 320</td>
<td>92</td>
<td>79</td>
<td>13</td>
</tr>
<tr>
<td>Aeronautical Sciences 490</td>
<td>47</td>
<td>32</td>
<td>15</td>
</tr>
<tr>
<td>Management 411</td>
<td>77</td>
<td>31</td>
<td>46</td>
</tr>
</tbody>
</table>

Assessment evaluators, who are also faculty familiar with the courses, engaged in a calibration or norming session of sample submissions using the University Ignite Rubric. This norming session served to discuss the specific levels of competencies represented in the rubric and ensure inter-rater reliability. After the norming session, the artifact samples were scored by evaluators without the knowledge of the grade the class section instructor had...
assigned to the artifact. The assessment scores and modality of instruction were collected and merged with the grade data for the artifact. Equating instructor-awarded grades with the assessment evaluation score was more of a subjective estimate than a one to one. The University Ignite Research Rubric had four levels of proficiency: Mastery, Practicing, Introductory, and Novice. Evaluators assigned a numerical score for each criteria level in the rubric: Mastery = 3; Practicing = 2; Introductory = 1; and Novice = 0. Prior to this exercise, faculty had chosen assignments with the targeted level of proficiency in mind.

Therefore, it would be expected that if a student scored a 90-100% on a mastery level assignment, the independent assessment evaluators would score the student’s performance at a mastery level. If the instructor’s score was perhaps a 95%, and the assessment evaluator scored the artifact at a 1 or a 2, then there was determined to be a mismatch in the scoring. In general, for comparison of the artifact’s course grade with the rubrics evaluation score, we assumed that if the course grade was between 90 and 100, then the average of the evaluation scores according to the rubric should be at the competency level of the assignment as identified by the faculty. For example, the average of a mastery level assignment should be expected to score almost a three using the rubric, and in between 90-100 as a course grade. The average of a practicing level assignment should be expected to score between a two and three using the rubric and in between 90-100 as the course grade. The difference between the two represented the competency level of the assignment.

To determine whether there are differences between course delivery modalities in both the course awarded grades and evaluator scores, a two-sample, two-tailed t-test assuming unequal variances was used for the sample from each course. To determine whether the course awarded grades and secondary evaluation awarded scores are higher in a traditional course delivery modality than the online course delivery modalities a two-sample, one-tail, t-test assuming unequal variances was used for the sample from each course.

With the artifacts, course grade and secondary evaluation data collected, for each course the following research questions could be answered:

Research Question 1: Does the artifact course awarded grade score equate to the secondary evaluation rubric score? Course grade = Secondary evaluation rubric

Research Question 2.1: Are the artifact course awarded grade scores equivalent between online and traditional delivery of curriculum? Online Course grade = Traditional Course grade.

Research Question 2.2: Are the secondary evaluation score equivalent between online and traditional delivery of curriculum? Online Secondary Evaluation Score = Traditional Secondary evaluation score.

Research Question 3: Are the online course awarded grades lower than traditional course awarded grades (because the rubric is being used correctly)?

4. Results

Research Question 1: Does the artifact course awarded grade score equate to the secondary evaluation rubric score? Using the entire sample of student artifacts collected, comparing the course awarded grade scores with the secondary evaluation rubric scores indicated that for all courses, there was a discrepancy in the proficiency level the student achieved according to the course instructors and the secondary evaluators. In all courses, the secondary evaluators were more conservative in their scores as shown in Table 2.

Table 2: Comparison of course awarded grades and secondary evaluation of student artifacts from academic year 2014-2015 used in the Ignite assessment activities

<table>
<thead>
<tr>
<th>Course</th>
<th>Competency Level</th>
<th>Course Awarded Average Score</th>
<th>Rubric Evaluated Score</th>
<th>Comparison</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Management 391</td>
<td>Practicing</td>
<td>92.61</td>
<td>1.20</td>
<td>Mismatch</td>
<td>Secondary evaluation did not confirm practicing level proficiency</td>
</tr>
<tr>
<td>Safety 320</td>
<td>Practicing</td>
<td>92.26</td>
<td>1.68</td>
<td>Mismatch</td>
<td>Secondary evaluation did not confirm practicing level proficiency</td>
</tr>
<tr>
<td>Aeronautical Sciences 490</td>
<td>Mastery</td>
<td>83.13</td>
<td>2.67</td>
<td>Mismatch</td>
<td>Secondary evaluation did not confirm mastery level proficiency</td>
</tr>
<tr>
<td>Management 411</td>
<td>Mastery</td>
<td>91.26</td>
<td>1.30</td>
<td>Mismatch</td>
<td>Secondary evaluation did not confirm mastery level proficiency</td>
</tr>
</tbody>
</table>
Research Question 2.1: Are the artifact course awarded grade scores equivalent between online and traditional delivery of curriculum?

Table 3: Comparison of course awarded grades between fully asynchronous and synchronous sections of student artifacts from academic year 2014-2015 used in the Ignite assessment activities

<table>
<thead>
<tr>
<th>Course</th>
<th>Competency Level</th>
<th>Course Awarded Average Score (asynchronous)</th>
<th>Course Awarded Average Score (synchronous)</th>
<th>t-Stat</th>
<th>P(T&lt;=t) two-tail test</th>
<th>Significant at the 95% level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Management 391</td>
<td>Practicing</td>
<td>92.17</td>
<td>94.09</td>
<td>-1.06</td>
<td>0.309</td>
<td>No</td>
</tr>
<tr>
<td>Safety 320</td>
<td>Practicing</td>
<td>93.39</td>
<td>85.50</td>
<td>1.94</td>
<td>0.100</td>
<td>No</td>
</tr>
<tr>
<td>Aeronautical Sciences 490</td>
<td>Mastery</td>
<td>87.78</td>
<td>72.44</td>
<td>2.51</td>
<td>0.003</td>
<td>Yes</td>
</tr>
<tr>
<td>Management 411</td>
<td>Mastery</td>
<td>98.09</td>
<td>86.67</td>
<td>5.99</td>
<td>2.41E-07</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Research Question 2.2: Are the secondary evaluation scores equivalent between online and traditional delivery of curriculum?

Table 4: Comparison of secondary evaluation of course awarded grades between fully asynchronous and synchronous sections of student artifacts from academic year 2014-2015 used in the Ignite assessment activities

<table>
<thead>
<tr>
<th>Course</th>
<th>Competency Level</th>
<th>Course Awarded Average Score (asynchronous)</th>
<th>Course Awarded Average Score (synchronous)</th>
<th>t-Stat</th>
<th>P(T&lt;=t) two-tail test</th>
<th>Significant at the 95% level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Management 391</td>
<td>Practicing</td>
<td>1.29</td>
<td>0.79</td>
<td>-3.77</td>
<td>0.001</td>
<td>Yes</td>
</tr>
<tr>
<td>Safety 320</td>
<td>Practicing</td>
<td>1.68</td>
<td>1.74</td>
<td>0.45</td>
<td>0.657</td>
<td>No</td>
</tr>
<tr>
<td>Aeronautical Sciences 490</td>
<td>Mastery</td>
<td>2.71</td>
<td>2.59</td>
<td>-0.79</td>
<td>0.439</td>
<td>No</td>
</tr>
<tr>
<td>Management 411</td>
<td>Mastery</td>
<td>1.17</td>
<td>1.50</td>
<td>3.31</td>
<td>0.001</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Research Question 3: Are the online course awarded grades lower than traditional classroom?

Table 5: Comparison of course-awarded grades between fully asynchronous and synchronous sections of student artifacts from academic year 2014-2015 used in the Ignite assessment activities

<table>
<thead>
<tr>
<th>Course</th>
<th>Competency Level</th>
<th>Course Awarded Average Score (asynchronous)</th>
<th>Course Awarded Average Score (synchronous)</th>
<th>t-Stat</th>
<th>P(T&lt;=t) one-tail test</th>
<th>Significant at the 95% level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Management 391</td>
<td>Practicing</td>
<td>92.17</td>
<td>94.09</td>
<td>-1.06</td>
<td>0.15</td>
<td>No</td>
</tr>
<tr>
<td>Safety 320</td>
<td>Practicing</td>
<td>93.39</td>
<td>85.50</td>
<td>1.94</td>
<td>0.05</td>
<td>No</td>
</tr>
<tr>
<td>Aeronautical Sciences 490</td>
<td>Mastery</td>
<td>87.78</td>
<td>72.44</td>
<td>2.51</td>
<td>0.012</td>
<td>Yes</td>
</tr>
<tr>
<td>Management 411</td>
<td>Mastery</td>
<td>98.09</td>
<td>86.67</td>
<td>5.99</td>
<td>1.21E-07</td>
<td>Yes</td>
</tr>
</tbody>
</table>

5. Discussion

5.1 Research Question 1

When examining the implications of the attempt to answer Research Question 1 (Does the artifact course awarded grade score equate to the secondary evaluation rubric score?), the authors determined that their results indicate that there is work to be done to ensure consistency. The secondary evaluators were less confident that students achieved the level of proficiency that was indicated by the instructor of the course. In every case, the application of a standard rubric did not achieve the desired results of consistency.

There are several possibilities that could lead to such a result. Discrepancies could be a result of the quality of the embedded rubric in the course and its ability to adequately measure the course learning outcomes. Because all rubrics were developed by subject-matter experts and were reviewed by colleagues, instructional designers and administration, it is highly probable that rubrics are effectively measuring course learning outcomes. Despite the unlikelihood of rubric quality as a major cause of discrepancies, it is important for course developers periodically to review all rubrics in the context of the learning outcomes covered.

A more likely scenario is the possibility that rubrics are being used inconsistently. Because secondary evaluators are “normed” in the use of their rubrics, it would be most useful to examine the use of rubrics by primary instructors (the instructors in the courses themselves).
Ideally, establishing training of primary evaluators (the actual instructors of the courses) in applying rubrics in a consistent manner will become a priority. Norming sessions held by each department can help to ensure that instructors are approaching the grading process similarly. Such training will also ensure the instructors understand the competency levels of the rubric.

Finally, instructors should be encouraged to approach rubrics not as a quick grading tool but as a way to clearly show students where they adhered to, or departed from, assignment guidance and to indicate where further work might be necessary to meet course learning outcomes. Training also will need to include a follow-up phase where the same type of study presented above is followed to see if the intended goal of course awarded grades equate to secondary evaluation of the rubric score.

5.2 Research Questions 2.1 and 2.2
Research Questions 2.1 and 2.2 examined the consistency of artifact scores between modalities. Specifically, 2.1 examined potential equivalencies in the course grades (Are the artifact course awarded grade scores equivalent between online and traditional delivery of curriculum?) while Research Question 2.2 focused on secondary evaluation of artifacts in synchronous and asynchronous modalities (Are the secondary evaluation scores equivalent between online and traditional delivery of curriculum?).

The results show that the use of consistent assignments, multi-modality templates and embedded rubrics were not successful in ensuring course section consistency across delivery modalities. As discussed above, multi-modality templates for use in synchronous delivery were developed from existing online courses. Therefore, online instructors were likely more familiar with the both assignment and the rubrics used in the course. Conversely, it is possible that both the assignments and their associated rubric are new to instructors in synchronous environments, accounting for grading discrepancies. Secondary evaluations confirmed the inconsistency. As with the conclusions from Research Question 1, a more likely scenario is the possibility that course content and rubrics are being used inconsistently. A solution to this inconsistency again appears to be establishing training (or “norming”) sessions for instructors. Additionally, specific MMT-focused training that carefully reviews the demands of the common assignments (as well as the criteria of their associated rubrics) could improve grading consistency across the modalities.

5.3 Research Question 3
The third research question looks at course grades in online and classroom sections of the same course (Research Question 3: Are the online course awarded grades lower than traditional classroom?). Clearly, the level of achievement of students is not consistently equivalent in asynchronous and synchronous offerings. There are several possible causes of this inconsistency. The first possibility could be discrepancies in the way that assignment is communicated to students. Online students may not have the benefit, for example, of verbal confirmation of assignment instructions. Additionally, in the case of common assignments, the synchronous instructor may not have the same familiarity with the assignments as the online instructors.

The authors could not cite specific reasons or trends for the differences in grades; however, there is a clear need for training, norming, mentoring sessions for instructors to ensure that students are receiving and gaining an equivalent learning experience across delivery modalities.

6. Conclusion
ERAU’s wide distribution of locations, accelerated term schedule, frequency with which courses are taught, and reliance on adjunct faculty have raised internal concerns about the consistency of the student learning experience and competency achievement. These concerns are also echoed at other universities and regional accrediting bodies in higher education. As a result, several steps have been taken to both create this consistency and to assess how well these efforts have fared indicate that, despite such intricately developed course content, instructions and embedded rubrics, consistency in achievement of learning outcomes, grading among instructors, and achievement across learning modalities has not been attained in the study sample. While this study only focused on a narrow group of courses, the researchers deem that the results are significant enough to recommend strategies to incorporate into the course management and mentorship of instructors to enhance the quality and equivalency of student learning across all modalities.

The researchers recommend periodic intensive training of instructors on assignments, skills, rigor of courses, and rubric meaning and uses.
Within the training, emphasis should be placed on the expected level of competency for assignments as measured by the criteria in common rubrics. Norming sessions using sample student artifacts and rubrics are of the utmost importance to “get everyone on the same page.” To address quality and equivalency of learning across various delivery modalities, the researchers recommend that instructors could benefit by teaching a course section in each modality to understand the nuances of each in an attempt to ensure consistency. Finally, it is important to close the loop of the course management process.

With structured course development and specialized training, the “lifecycle” of templated online courses, and the MMT assignments which are developed from them, should involve frequent review. This examination should include periodic use of secondary evaluations in assessments to ensure the consistency of student learning experience across delivery modalities and inform future curriculum content changes and instructor training requirements.

ERAU creates a climate where it encourages constant review to determine best practices. Since the study was concluded, Embry-Riddle Aeronautical University Worldwide has been ranked as the best Online Bachelor’s Program for two consecutive academic years in 2016 and 2017 (U.S. News & World Report L.P., 2017). This achievement has been greatly credited to the rigorous work performed by all stakeholders in the university, to enhance the students’ overall course experience, delivery mode, and instructor experience, according to the research criteria established in the IGNITE Quality Enhancement Plan.

References


