

Full Length Research Paper

A graduate education program's look at grading

Oluwatoyin Adenike Akinde

Department of Graduate Teacher Education, College of Education, Concordia University St. Paul, United States.

Received 16 July, 2019; Accepted 4 September, 2019

This study on grading was based on a mixed-method design. Assessment and evaluation have been used synonymously with grading, and have a deeply seated significance in the matters of curriculum, learning outcomes, and instructional strategies, to name a few. A point of convergence in the literature is that grade is an indication of student's mastery, that is, what student knows, understands and can do. In the process of programmatic assessment, the department of graduate education increased its grading scale. The before and after data, which was the student's grade, was evaluated to determine if the increase in grading scale (treatment) had an impact on the frequency of students who earned high scores in a particular course. The result shows that the treatment did not have an impact on the number of students that earned A. In a further analysis, it was evident that students in the graduate program advanced their effort and employed a greater level of rigor in addressing course assignments after the grading scale was increased. What was also established from this study is that the use of a rubric as a grading tool fostered transparency.

Key words: Grading, assessment, evaluation, rubric, adult learner.

INTRODUCTION

Grades are used by educators to show students' mastery. It summarizes students' achievement in a content area (Allen, 2005). Educators use a set of criteria to determine what students know, understand and can do (Heacox, 2009). There are two sides to the debate on grading. One side has argued that grades alone are not indicators of what students know and can do, while the other side contends that grades are indicative of what students understand, coupled with the teacher's evaluation of the students' work (Allen, 2005).

The traditional method for reporting students' grades includes the conversion of numeric scores into letter grades of A, B, C, D, and F, a practice that has been in existence in secondary education since the 19th century (Brookhart, 1994; Cox, 2011). A quantitative system of grading in higher education was traced back to the 18th century, credited to William Farish (Soh, 2010).. An A grade is usually awarded for excellent work, B is considered very good, C is considered average, D is considered less than average, and F is for

E-mail: akindefakuajo@csp.edu.

Author(s) agree that this article remain permanently open access under the terms of the [Creative Commons Attribution License 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

failing. There are several philosophies and variations around the grading of students' work and the criteria graded. Whereas some educators grading include criteria such as grammar, style, timeliness, and neatness in addition to the quality of the material produced (Guskey and Pollio, n.d.), other educators may be more interested in the students' creativity, that is, the quality of work produced. The difference is that one educator is more interested in the whole student and sees the importance of rewarding several efforts leading to the final product whereas the other educator places emphasis simply on the finished product. This is a means to an end versus an end in itself approach.

The terms *assessment* and *evaluation* are often used interchangeably in empirical matters relating to accreditation and discussion about how to support students' learning. According to Morrison et al. (2010), *assessment* is the systematic collection of data used to evaluate, that is, to judge the quality or success of a program or student. Grading, which falls under the umbrella of assessment, has a few typologies, including confirmative assessment, which is predicated under the principle of continuous evaluation and extends beyond summative evaluation.

However, the most commonly used practices are summative and formative assessments. As noted in Taras (2005), *summative assessment* is a teacher's judgment whereas *formative assessment* is summative assessment coupled with feedback from the teacher for the use of the learner. Advocates of summative assessments take into account a student's overall performance in a given course or activity whereas those favoring formative assessment zone in on the learning process and many activities leading to the student's outcome in the course (Popham, 2000). Clearly, the use of summative and formative assessments adds to the variations of grading because these two types of assessment are conceived differently from teacher to teacher, which further compounds the complexity of assessment (Taras, 2005).

Institutions have the liberty to determine what percent equals a specific letter grade. However, variations may exist among educators in the same school in the grading percentage and criteria used to determine a letter grade. For instance, some consider 90 to 100% as meriting a grade of A (Guskey and Pollio, n.d.) as opposed to 96 to 100% being an A. While grades represent students' mastery, grading rubrics are used in many institutions of higher education to promote transparency in the assessment of students' work. Rubrics, which are regarded as tools for assessing students work, are becoming commonplace in educational systems. Proponents of such tools favor their use because they clearly explain the criteria for what constitutes *excellent*, *good*, *average*, and so on. In many instances, rubrics

include percentage; provide a delineation of what is considered adequate, developing, and below expectation; and includes some quantification about the frequency of certain tasks, such as the inclusion of three bibliographies as opposed to two. The use of rubrics as a grading tool seems to reduce some degree of subjectivity in the assessment of students' work (Andrade, 2000). Andrade's (2000) position is that rubrics can be used as an instructional tool because they make teachers' expectation clear for the teacher and the students. Rubrics are beneficial for teachers because they have the potential to help the teacher assess student's mastery and command of the content, which can also be linked to the course learning outcome. A teacher assessing the quality of a student's work makes a judgment about what is deemed as high versus marginal quality; whereas rubrics support an efficient and educated assessment. Reddy and Andrade (2010) also reported that in most cases when rubrics are presented to students and used in assessing students' work, most students, especially adult learners, tend to excel and produce high-quality work.

On the premise of variation in the grading scale, some institutions support the use of extra credit for activities that are above and beyond expectation whereas some do not. In Guskey and Pollio (n.d.), some schools espouse a binary approach in grading students' work, using P for pass and F for fail. Some institutions do not report a grade of F for students. The term *No credit* or *Incomplete* may be used for work below passing—that is, in lieu of fail. Undoubtedly, grading is surrounded by many more controversies.

Grade inflation is not a new phenomenon in higher education. Students, including those in graduate programs, are required to demonstrate their learning in various forms, including writing a paper, contributing to a group project, conducting a case study, providing reflections, carrying out research, developing a portfolio and/or other creative activities, to name a few. Test-taking, although frequently used to evaluate students' mastery of the content in undergraduate programs, is rarely administered for graduate students.

This paper focuses on how a graduate program in a Midwestern U.S. university has attempted to approach the issue of grade inflation in the education programs. Its faculty member and administrators considered what it meant for students to earn an A or B. A discussion emanated from developing a rubric to accurately reflect a higher expectation for an A grade. Part of the discussion was how to increase the rigor of work that was worthy of a grade of A as opposed to a grade of B.

In this research, the terms *core course*, *grade inflation*, *grading scale*, and *graduate education program* are used. The author recognizes that these terms could be understood differently from one setting to another. Therefore, the terms are operationalized as appropriate

Table 1. Old versus new grading scale of institution under study.

Grade	Old grading scale		New grading scale	
	Point range	Percent	Point range	Percent
A	675-628	100-93	710-682	100-96
B	627-567	92-84	681-611	95-86
C	566-506	83-74	610-540	85-76
F	< 506	< 74	< 540	< 75

to the context of this study.

Core course

In the institution where this research took place, some course work is generalized to all graduate education programs. Core courses are required for all students in the master's degree in education program. Although the sequence of the core courses varies from one program to another, the learning objectives, content, text material, assignments, and rubric are the same across the programs, that is, a core course in one program can be administered in the first semester but can be in the second, third, or fourth semester in another program. This is likely a limitation because the level of writing and depth in analysis of a first-semester student in a graduate program is likely to be different for the same student after 1 year of the graduate learning experience. Consequently, when assessing students' material to determine if students increased the rigor of their writing and academic output in the program, the researcher of this work was considerate of the course sequence, that is, when students took the course and produced the material that was part of this study. The chosen core course that was used in assessing the quality of work produced by students was the Educational Research and Applications course, which is generalizable across the master of arts in education program in terms of learning objectives, outcomes, content, and course material. However, the point in which students take the course vary depending on the emphasis area. The said course landed between course number two and six in the graduate education program.

Grade inflation

Grade inflation occurs when students are readily awarded an A for less than excellent work and a B for mediocre work (Kohn, 2002). However, granting a grade of A for less than excellent work is not the only factor. Rather, what causes grade inflation is the higher frequency of As and Bs. Therefore, when everyone or the majority of the

class earns an A, the ranking of excellent has little to no value. The issue of grade inflation is considered rampant among many universities and colleges even though many institutions avoid the problem (Johnson, 2006). However, for a college interested in assessing its student output, the department under study has considered many ways to assess its instructional practices as well as increase the rigor of its program. Hence, it has looked at the percentage of students earning an A or B and implemented a higher grading scale.

Grading scale

A grading scale is the conversion of numeric scores into an alphabet grade. Higher institutions have various criteria for grades. In some colleges and universities, a 100 to 90% score equals a grade of A, 89 to 80% would be a grade of B, whereas C is 79 to 70% (Creswell, 2002). Some schools allow extra credits; however, this is not a practice in the institution under study. The literature does not suggest that a grade of B in one school would equal a grade of B in another. However, in the case whereby students are transferring credits from one institution to another, most of the time, the final grade would not be transferred using the earned grade; rather, institutions are likely to use its system of transferring grades, such as grade point average. The graduate education program under study does not consider a grade of less than C as passing and does not award minus grades such as A. The institution under study uses the grading scale in its graduate education program reflected in Table 1, which provides a comparison of the institution's new, present scale with its previous, old scale.

Based on the espoused grading scale for the graduate programs, as expressed in the graduate handbook in the College of Education under study:

1. A grade of A indicates a superior level of understanding and expression of ideas, with a depth of critical thinking on issues such that the individual shows a profound level of understanding of the material.
2. A grade of B indicates that the student exhibited good

basic understanding and diligence, and was able to extend the knowledge to other situations, making connections between the material and other concepts.

3. A grade of C indicates that the student exhibited an acceptable basic understanding of the material and was able to express that understanding clearly and accurately. This shows a preprofessional level of understanding.

4. A grade lower than C indicates that the student exhibited a lack of basic knowledge and understanding of the material.

It is important to note that one of the charges of the graduate policy committee of the institution under study, which is comprised of faculty and staff, is to determine grading policies, including grading scale. Faculty and administrators of each college and department may suggest and present a grading scale to the graduate committee. Before the implementation of the higher grading scale in the graduate education program, the change was proposed to the dean of the college and tendered to the graduate policy committee.

Graduate education program

A graduate education program looks different across many institutions. Some are initial licensure programs whereas some are not. In both cases, students have earned an undergraduate degree prior to entering the graduate program. Students enrolled in the graduate program enter the program for career advancement and to further develop their craft as professional educators.

LITERATURE REVIEW

The higher education community sees graduate education as a means for career advancement and, in some cases, career preparation; and as a way to increase the depth of professionals already in their line of work. The practice of teaching in higher education varies for those in undergraduate and graduate programs. In teaching for pre-K–16, in most cases, the knowledge resides with the teacher (Davis and Hoffman, 2008), whereas, for professional educators in graduate programs, teaching is the exchange of knowledge and exploration of information. Teachers in a pre-K–16 program identifies with the principles of pedagogy whereas those teaching in graduate programs tends to embody the principles of andragogy, heutagogy, and ergonagy in their work with students (Reynolds et al., 2009).

The debate about andragogy versus pedagogy revolves around teaching and learning, wrote Yonge (1985). Bloom's taxonomy (Appendix 1), is often used across the education system to determine the degree to which students as learners are able to remember,

understand, apply, analyze, evaluate, and create, respectively (Krathwohl, 2002). The first three levels of the learning taxonomy are associated with fundamental thinking skills, which is typified by the ability to recall, explain, and use information in a new situation. Whereas the top three levels are associated with an advanced level thinking skill, which includes drawing connections, appraising, and producing new information. This illustration is consistent with the graphical representation by Tolks et al. (Figure 1).

The first three from the bottom of the learning taxonomy include knowledge, comprehension, and application. In moving up in Bloom's taxonomy, when students analyze information, they are assessing the information and its use. When students evaluate, they are making a judgment as to the value of the information. *Create*, which is at the top of the learning taxonomy, implies synthesizing information—that is, taking a look at various information to make a judgment. Teaching adult learners in higher education using the principles of andragogy requires the use of high-order thinking, representing the top three layers of the learning taxonomy: Analyze, evaluate, and create (Wang and Farmer, 2008). Working with graduate students requires innovative instructional approaches including building on the prior experiences of students, which is not to say that teaching graduate students is incongruent with the first three levels of the learning taxonomy. Rather, the first three layers of the taxonomy are a starting place for the adult learner, which may be necessary for developing a frame of reference for engaging in scholarly discourse—inherent in teaching adult learners. In graduate programs, the first three layers of the learning taxonomy are typified by reading and interpretation of texts, which is a necessary first step for engaging in intellectual dialogue. Reading and review of academic material serve as a pretext to the greater intellectual experience that manifests when students as adult learners engage in cooperative learning or other forms of a learning exercise. The work of Heacox (2009) Tolks et al. (2016) and Krathwohl (2002) can be used to explain the intersection between the learning taxonomy and grading criteria. Student's grade, which is an indication of what the student know, understand and can do (Heacox). Knowing is an indication of remembering, understand is an indication of having a grasp of the information, and can do is the student's ability to apply the learning (Tolks; Krathwohl). The appendix section provides further illustration of this intersection, and how they serve as an overlay to the grading construct espoused in the department of graduate education under study.

Many accelerated learning programs use the flipped classroom model to facilitate learning among adult learners and promote self-directedness (Tolks et al., 2016). Under this tenet, higher educators tend to front-load the reading material, whereby the review of

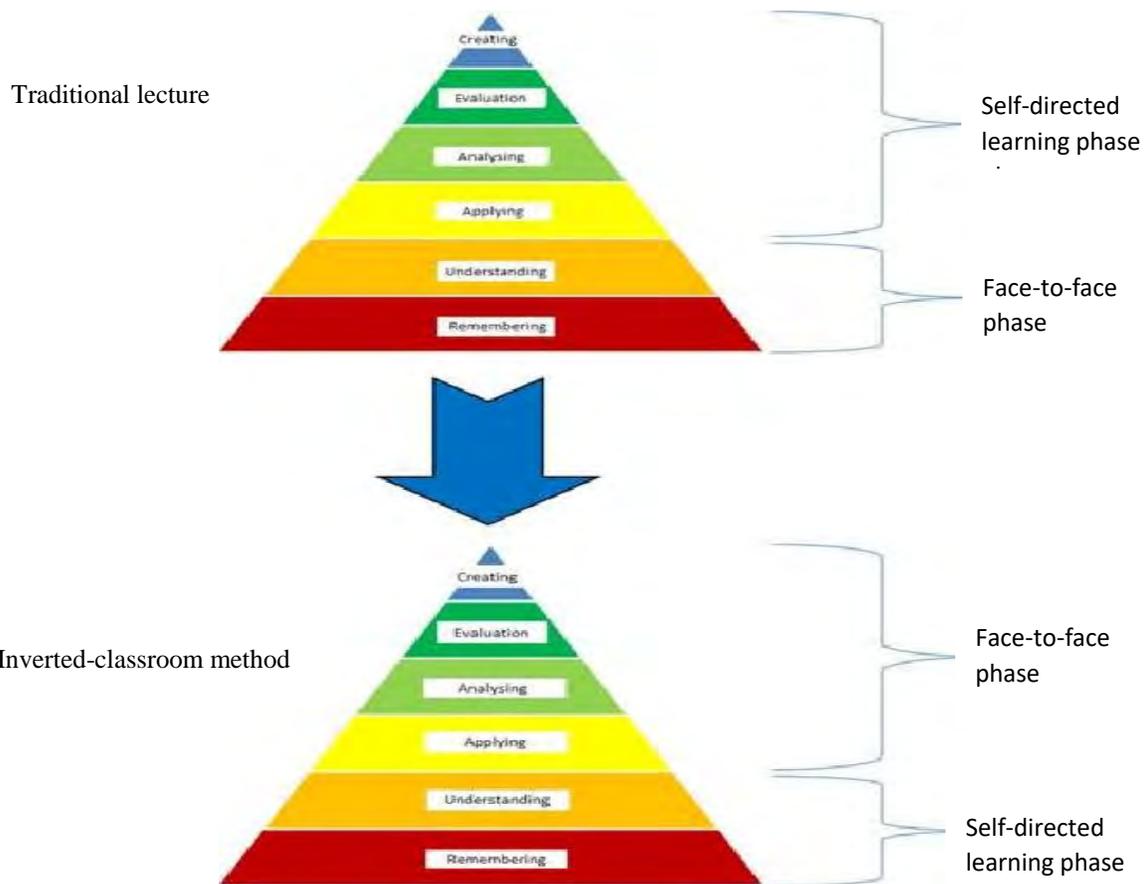


Figure 1. Blooms' taxonomy and the instructional method. From "An introduction to the inverted/flipped classroom model in education and advanced training in medicine and in the healthcare professions" by Tolks et al. (2016), *GMS Journal for Medical Education*, 33.

academic material helps accelerate learning and create a shared language to be used in the classroom among the professional learning community (Wlodkowski, 2003). This student-centered teaching, rooted in the philosophy of andragogy, is highly regarded for promoting higher-order learning because the teacher is able to efficiently move up the learning taxonomy in a manner that is stimulating for the learners. Tolks et al. (2016) graphical representation shows the connection between Bloom's taxonomy, the difference between the traditional versus the flipped classroom model (also known as the inverted-classroom method), and the self-directed learning level in the learning taxonomy.

Similarly, in a growth curve analysis of students in mathematics, Ahmed et al. (2013) postulated that several factors can change and impact students' learning and growth. Among other things, they reported that changes in positive emotions are systematically associated with changes in self-regulated learning and achievement. This is consistent with Edwards (2004), who reported that

instructional approach matters significantly to students' learning.

METHODOLOGY

Research questions

1. Did increasing the percentage to earn an A grade reduce grade inflation in the graduate education program?
2. Was there a difference in the number of students who did not earn an A?
3. Based on comparing students' work before and after the use of the higher grading scale, what phenomenon was observed about the quality of work produced?

Research design and data collection

This mixed-method research drew on the foundations of positivist and postpositivist paradigms. The first espoused methodology included a quantitative approach for evaluating the final grade of students in the graduate education program over the period of one

academic year before and after the implementation of the increased percentage to earn an A. The data, final grades, were compared from before and after for students who did not earn an A. The final grades, although not compared to the previous year on an individual student basis, were compared with data from a year prior to the implementation of the higher scale to determine if the treatment had an effect on students' grades and, if it did have an effect, the degree to which it had an impact.

In the second methodology, students' papers, which were randomly selected, were accessed through the course management system. This plan was communicated with the department chairperson and college dean. A qualitative approach was used to evaluate the students' papers to determine if students were more rigorous in completing graded assignments in an effort to earn an A. Notably, the said course is used in all the Master of Arts in education program emphases areas. The scoring rubrics and papers compared for the two academic years, 2016 (pre) and 2017 (post) in an effort to assess the quality of students' work before and after the higher grading scale was implemented. The students' papers that were reviewed came from each of the emphases areas in the graduate education program. The random selection meant that a list with students' names was made and one name was selected randomly from the list. Therefore, any of the students' paper from the course had an equal chance of being selected for the review. This randomization method was used for the before and after groups for the six emphases, which generated 12 papers. Once the randomization process was completed, the selected papers were printed and the students' information was redacted. Then, the papers were labeled *before* and *after*. Although a randomized protocol was applied to the selection of papers included in the qualitative approach, the volume of materials generated the and the number of students' papers included were not enough to suggest that the evidence was representative of the other students who also earned an A but not included in the review.

Table 2 provides an illustration of the research question, the data used to assess the question, and the procedure that was used to collect and analyze the results.

Population and samples

The graduate education program comprised of 77.4% women, and 22.6% men in 2016, and 77.7% women and 22.3% men in the 2017 academic year, who have earned their undergraduate degree and have work experience prior to entering the program and, therefore, are adult learners. The data that were used for this study was the final grades for graduate students in the education program—that is, students in the master of arts in education program. The data included students' final grades in the curriculum and instruction, differentiated instruction, early childhood education, educational leadership, educational technology, and special education programs in the College of Education. The program is a 2-year program leading to a master's degree in education with an emphasis in any of the aforementioned specializations. The students in the program took a set of core courses, which ranged from 30 to 50% of the course work, whereas the other courses are content area specific. The early childhood program is unique because students in that specialty take one additional course to fulfill the requirements of the program.

Although the duration of the program is 2 years, comprising of courses and capstone research, some students may take a longer time to complete the program as some take a break for a period of time for professional, personal, and other reasons. Additionally, it should be noted that students are able to take the course in two modalities: The traditional on-campus, face-to-face or the online format. The course requirements and learning activities are similar

for students in either format.

In assessing if the increase in the percentage to earn an A grade reduced grade inflation, chi-square was used to determine whether the implementation of the higher grading scale changed the number of students who earned a final grade of A based on comparing before- and after-data. Also, a *t*-test was used to determine the difference in before- and after-data. Before the implementation of the higher grading scale, students were required to have 93 to 100% to earn an A, whereas, for the new scale, students needed 96 to 100% for an A. Those who did not earn an A were those who scored 95% and below. The students' grade that was part of this assessment included all students who earned an A.

The 12 randomly selected students' work, that is, final papers across the graduate education programs were reviewed, along with the grading rubric. The students' work that was reviewed comprised of six papers written before and six papers written after the increase in grading percentage went into effect. Although this was not intended to be representative of the student body, it was intended that this review would provide additional insights for addressing the research question and, more importantly, the quality of work that received an A. Although tangential to the espoused research design, the observation from the artifact review (students' papers) was corroborated with observations of four faculty members in the graduate education program, who are responsible for teaching and assessment in the various emphases areas.

Privacy and consent

The researcher of this study contacted the institutional research office of the institution under study via e-mail, explaining that the researcher, upon recommendation of colleagues in the Department of Graduate Teacher Education, wished to conduct an evaluation of the number of students earning an A before (2016 grades) and after (2017 grades) the implementation of a higher grading scale, taking into account the entire student body during the specified academic years. The Office of Institutional Research at the university responded with a statement of support and agreed to provide the necessary data.

The significance of the study

The primary significance of this study was to help the faculty and department evaluate its practice. As such, this study shows that the implementation of a higher grading scale led to increased rigor, which also changed the narrative on grade inflation. In addition, this study serves as implications to other graduate programs in increasing the quality of students' academic output while reducing grade inflation. Thus, this research adds to the body of knowledge.

RESULTS

The statistical software that was used for the data analysis was Minitab 16. Cross-tabulation and chi-square were used to determine whether the change in the grading scale (before 93% = A versus after 96% = A) had an impact on the degree to which students earned an A, and the *t*-test was used to determine where there was a difference in the categories of A, Not A, IP, and Pass. In this report, the implementation of the higher grading scale to earn an A was considered a treatment. Therefore, the before- and after-data of students who earned an A were

Table 2. Research framework.

Research question	Method of assessment	Procedure
1. Did increasing the percentage to earn an A grade reduce grade inflation in the graduate education program?	Quantitative: Situated in the positivist paradigm, the researcher took an objective view in analyzing students' data using chi-square to determine the degree to which there was a difference in the number of students who earned an A before and after the implementation of the higher grading scale. The result of this analysis was used to determine if the change in grading scale had an impact	The percentage of students who earned an A was compared to those who did not earn an A before and after the implementation of the higher grading scale.
2. Was there a difference in the number of students who did not earn an A?	A <i>t</i> -test was used to assess grades before and after implementation of the higher grading scale to determine the difference in the categories of those who earned an A and those who did not earn an A	
3. Based on comparing students' work before and after the use of the higher grading scale, what phenomenon was observed about the quality of work produced?	Qualitative: In an effort to address Research Question 3, the data collection and analysis resided in the postpositivist framework. Thus, the students' papers were treated as artifacts and were examined for rigor and quality. The artifacts were comprised of students' final papers before and after the implementation of the higher grading scale.	Students' final papers in the aforementioned course were reviewed in an effort to determine if students increased their effort to meet the rigor to earn an A grade after the implementation of a higher grading scale. A comparison was made using students' material from the same course before the use of the higher grading scale.

Table 3. Chi-square test results.

	Grade				Total
	A	Not A	In progress	Pass	
Before	1657	93	12	14	1776
	1598.06	161.56	7.91	8.47	1598.06
	2.174	29.092	2.117	3.605	2.174
After	1172	193	2	1	1368
	1230.94	124.44	6.09	6.53	1230.94
	2.822	37.769	2.748	4.680	2.822
Total	2829	286	14	15	3144

Chi square = 85.007, *df* = 3, *p*-value = 0.000.

compared in an effort to address the research question (Tables 3 and 4).

The chi-square does not distinguish between the four categories (A, Not A, In Progress, Pass). Based on the *p*-value of 0.00, as shown in Table 3, which is smaller than 0.05, the treatment, in this case, the implementation of a higher grading scale had an impact in at least one of the categories.

The *t*-test was used to examine the individual category to determine if the A category was affected by the

treatment. When the *t*-test was used to compare the difference between how many students earned an A before and after the treatment, it appears that based on the *p*-value of 0.396 being larger than 0.05, there is no evidence that there was a significant difference in before and after the implementation of the higher grading scale for those students who earned an A grade (Table 4).

From the *t*-test (Table 5), the *p*-value of 0.006 being smaller than 0.05, there is some evidence that the treatment had an impact on students who did not earn an A.

Table 4. Two-sample T-test and confidence interval data: A Grade.

	N	Mean	Std dev	SE mean
After	3	914	380	220
Before	3	1286	532	307
Difference = mu (after) - mu (before)				
Estimate for difference: 373				
95% confidence interval for difference: (-1574, 829)				
T test of difference = 0 (versus. not =): <i>t</i> -value = -0.99, <i>p</i> -value, = 0.396, <i>df</i> = 3				

Table 5. Two-sample T-test and confidence interval data: Not A Grade.

	N	Mean	Std dev	SE mean
After	3	182.3	15.1	8.7
Before	3	90.3	17.2	9.9
Difference = mu (after) - mu (before)				
Estimate for difference: 92.0				
95% confidence interval for difference: (50.0, 134.0)				
T test of difference = 0 (versus not =): <i>t</i> -value = 6.96, <i>p</i> -value = 0.006, <i>df</i> = 3				

The box plots in Figure 2 provide insight, whereas the *t*-test shows the statistical significance or the lack of it. Therefore, the *t*-test is more precise. In Figure 2, the lines in the box plots represent the median and the dots represent the means. The line across both after and before categories represents the difference between the means. Visually, using Figure 2, because of the overlap in the category of students who earned an A, there is likely no significant difference in before and after whereas in the Not A category, there was no overlap.

It appears that after the treatment, there were fewer students who earned an A when comparing the mean from Tables 4 and 5, which shows that before the implementation of the higher grading scale, there were 1,289 students who earned an A compared to 914 students who earned an A after the treatment. Students who did not earn an A before were 90 compared to 182 students after the treatment.

DISCUSSION

The difference in the number of students who earned an A was not significantly different, based on the comparison of data on students’ final grade before and after the implementation of a higher percentage. Following are the study findings by research question:

1. Did increasing the percentage to earn an A grade reduce grade inflation in the graduate education program? No, based on the *p*-value of 0.396, there is no significant

difference in before and after the implementation of the higher grading scale for those students who earned an A grade.

2. Was there a difference in the number of students who did not earn an A? Yes, based on the *p*-value of 0.006, there is some evidence that the treatment had an impact on students who did not earn an A.

3. Based on comparing students’ work before and after the use of the higher grading scale, what phenomenon was observed about the quality of work produced? In comparing the students’ selected papers for before and after the increased grading scale, there was some evidence to suggest that students increased their efforts to earn a higher grade. In almost all instances after the increased grading scale, students were particularly thorough in addressing the rubric criteria.

The students’ work reflect a particular level of quality; however, those papers produced by students after the treatment was of a much higher level of graduate writing in that the level of depth, analysis, and reflection as presented in the papers was undoubtedly rigorous. Notable, the inclusion of a rubric made a clear difference in whether students addressed the criteria for the assignment and the degree to which they did. In almost all instances, students provided the illustrations necessary yet concisely addressed the requirements.

Evidently, after the higher grading scale, students were even more incisive in addressing the criteria for the assignment.

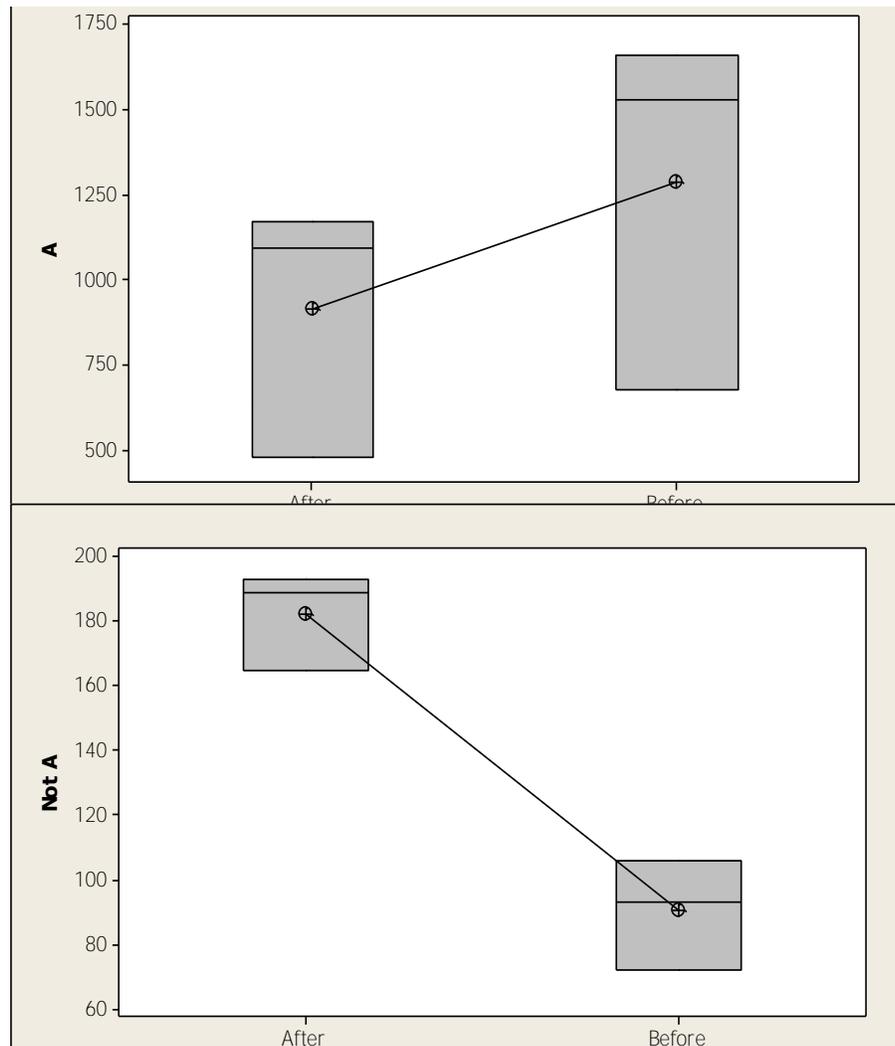


Figure 2. Comparison of before and after implementation of higher grading scale for A grades and not-A grades.

Conclusion

To circle back to Reddy and Andrade (2010) on the use of rubrics, it is likely that students who earned an A were able to excel to the new standard of grading. Because the criteria were openly laid out and communicated, students sampled in this study were able to meet the requirements for academic excellence. Although the number of students who excel and earn good grades is considerably higher, grade inflation remains perpetual, which is likely an indication that educators and institutions of higher education could start looking at grade inflation differently. When rubrics are used to grade students on specific criteria, it demystifies the ambiguity about learning objectives, teacher expectation, and students’ outcome. Therefore, students can demonstrate what they know,

understand, and can do. Perhaps grade inflation, which is caused by a majority of students’ earning academic honor, can be attributed to students’ motivation to do well.

Limitations

It was not clear if there was a difference in the instructional approach used in instructing students before and after the treatment. That is, it was not determined if the instructional efforts were intensified after the treatment. Because no such data were accounted for, this is a limitation to the results. Although there is valuable information derived from this study that could serve as an implication to other institutions of higher

education, this finding can only be generalized to the department where the research took place, which means other institutions of higher education could consider this research as a frame of reference; however, they have the opportunity to replicate this study in their institution.

Recommendations for future research

What was not explored in this study is a measure of a student's grade in relation to persistent enrollment. Thus, future research is necessary to determine if students' learning increased over time after persistent enrollment in the graduate program. Furthermore, the criteria for measuring students learning overtime would need to be determined for this assessment to occur.

It was gathered from this study, though a hypothesis at best, that students are likely motivated and invigorated by the new expectation required to earn an A. Future study is required to examine this further and to explore the potential Hawthorne effect that would likely result from the change in the grading scale.

There are other factors that are considered an outlier to this study. One of these factors is the evaluator's judgment when assessing students' work and awarding grades. A future researcher could consider the evaluator's shift in paradigm when using a new grading scale to assess students' work.

In the matter of reducing grade inflation, this researcher wonders if some schools, such as highly competitive schools, naturally attract students who are likely to earn a high grade based on having a high expectation for quality scholastic work. To that end, could this be used to explain grade inflation? A comparison between the grading data of students in private and public institutions is also needed.

CONFLICT OF INTERESTS

The author has not declared any conflict of interests.

ACKNOWLEDGEMENT

The author would like to express her appreciation to the Dean of the College of Education and Chair of the Department where this study took place, for their support. She is also grateful to colleagues who were part of the work leading to this research.

REFERENCES

Ahmed W, Van der Werf G, Kuyper H, Minnaert A (2013). Emotions, self-regulated learning, and achievement in mathematics: A growth curve analysis. *Journal of Educational Psychology* 105(1):150-161.

- doi:10.1037/a0030160
- Allen JD (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House* 78(5):218-223. doi:10.3200/tchs.78.5.218-223
- Andrade HG (2000). Using rubrics to promote thinking and learning. *Educational Leadership* 57(5):13-18. Retrieved from <http://www.ascd.org/publications/educational-leadership/>
- Brookhart SM (1994). Teachers' grading: Practice and theory. *Applied measurement in Education* 7(4):279-301.
- Cox KB (2011). Putting classroom grading on the table: A reform in progress. *American Secondary Education* 40(1):67-87. Retrieved from ERIC database. (EJ951229)
- Creswell JW (2002). *Educational research: Planning, conducting, and evaluating quantitative*. Upper Saddle River, NJ: Prentice Hall. pp. 146-166.
- Davis RP, Hoffman JL (2008). Higher education and the P-16 movement: What is to be done? *Thought and Action*. Pp.123-134. Retrieved from https://www.nea.org/assets/img/PubThoughtAndAction/TAA_08_15.pdf
- Edwards MC (2004). Cognitive learning, student achievement, and instructional approach in secondary agricultural education: A review of literature with implications for future research. *Journal of Vocational Education Research* 29(3):225-244. doi:10.5328/jver29.3.225
- Guskey TR, Pollio HR (n.d.). Grading systems: School, higher education. Retrieved from <http://education.stateuniversity.com/pages/2017/Grading-Systems.html>
- Heacox D (2009). *Making differentiation a habit: How to ensure success in academically diverse classrooms*. Minneapolis, MN: Free Spirit.
- Johnson VE (2006). *Grade inflation: A crisis in college education*. New York, NY: Springer Science and Business Media.
- Kohn A (2002). The dangerous myth of grade inflation. *Chronicle of Higher Education* 49(11):B7-B9. Retrieved from ERIC database. (EJ657735)
- Krathwohl DR (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice* 41(4):212-218. doi:10.1207/s15430421tip4104_2
- Morrison GR, Ross SM, Kemp JE, Kalman H (2010). *Designing effective instruction*. Hoboken, NJ: Wiley.
- Popham WJ (2000). *Modern educational measurement: Practical guidelines for educational leaders*. Boston, MA: Allyn and Bacon.
- Reddy Y, Andrade H (2010). A review of rubric use in higher education. *Assessment and Evaluation in Higher Education*, 35(4):435-448. doi:10.1080/02602930902862859
- Reynolds J, Laton D, Davis T, Stringer D (2009). From pedagogy to heutagogy: A teaching-learning continuum. *Florida Community College at Jacksonville*.
- Soh KC (2010). Grade point average: what's wrong and what's the alternative? *Journal of Higher Education Policy and Management* 33(1):27-36.
- Taras M (2005). Assessment—summative and formative – Some theoretical reflections. *British Journal of Educational Studies* 53(4):466-478. doi:10.1111/j.1467-8527.2005.00307.x
- Tolks D, Schäfer C, Raupach T, Kruse L, Sarikas A, Gerhardt-Szép S, Sostmann K (2016). An introduction to the inverted/flipped classroom model in education and advanced training in medicine and in the healthcare professions. *GMS Journal for Medical Education* 33(3):Doc. 46. doi:10.3205/zma001045
- Wang V, Farmer L (2008). Adult teaching methods in China and Bloom's taxonomy. *International Journal for the Scholarship of Teaching and Learning* 2(2), Article 13. doi:10.20429/ijsotl.2008.020213
- Wlodkowski RJ (2003). Accelerated learning in colleges and universities. *New Directions for Adult and Continuing Education* 2003(97):5-16. doi:10.1002/ace.84
- Yonge GD (1985). Andragogy and pedagogy: Two ways of accompaniment. *Adult Education Quarterly* 35(3):160-167. doi:10.1177/0001848185035003005

Appendix 1

Appendix 1. Grading construct in relation to assessment and bloom's taxonomy.

Heacox (2009)	Tolks et al (2016) on bloom's taxonomy	Department of graduate education grading criteria
Know	Level 1: Remember, that is to recall the information	A grade of C- indicates that the student exhibited an acceptable basic understanding of the material and was able to express that understanding clearly and accurately – Shows a preprofessional level of understanding
Understand	Level 2: Understand-to have a grasp of, that is, to recognize and be able to explain the information	This is also consistent with the criteria to earn a grade of C, meaning student in this category is expected to be able to demonstrate and express understanding of the course content clearly at the preprofessional level
Can do	Level 3: Application, which is to use the information in new setting.	A grade of B indicates that the student exhibited good understanding and diligence, and was able to extend the knowledge to other situations, making connections between the material and other concepts. The expression of these ideas shows a greater depth of understanding and critical thinking
	Level 4: Analyze, which is to draw connection, organize and examine. Level 5: Evaluate, to justify and appraise. Level 6: Create, produce, design, construct, develop, formulate.	A grade of A indicates a superior level of understanding and expression of ideas, with a depth of critical thinking such that the individual shows a profound level of understanding. Critical thinking, as typically expressed in the graduate rubric, the individual is required to analyze and draw inferences with exceptional clarity. Students in this category are expected to demonstrate a greater level of understanding, by formulating questions or response that reflects a particular depth of scholarship, which includes assessment, and the creation of new information-contributing to the body of knowledge both theoretically and pragmatically