On The Experience
Of The Arab Open University
With The Standard Deviation Method

Abdulkarim S. Al-Eisa, Arab Open University
Abdulla M. Alhemoud, American University of Sharjah, UAE

ABSTRACT

The Arab Open University (AOU) has adopted the standard deviation method (SDM) as a grading system in replacement of a fixed scale. Adopting SDM was intended to remedy a problem that has resulted from discrepancies between AOU's graduation requirements and those of its partner, UK Open University. This paper aims mainly at investigating whether SDM has served the purpose for which it was selected. A data set of the final letter grades of students enrolled in 18 licensed courses from UK-OU at Kuwait branch was used. These letter grades were analyzed in comparison with the letter grades that would have been assigned to students had the fixed grading scale been utilized. The results of the comparative analysis revealed that SDM resulted in benefiting 39.4% of all students and lowering the letter grades of 10.4% of the students. Despite its positive results, SDM has not contributed significantly to remedying the problem in question. Thus, a case for withholding the D grade while continuing with using SDM was presented.

Keywords: standard deviation method, student grading, Arab Open University.

INTRODUCTION

What to do about grading or how to do it is often less a matter of correctness and more a matter of preference and perceived value or importance.

David A Frisbie and Kristie K. Waltman, 1992

*I*naugurated in June 2001 as a Pan Arab project under the auspice of the Arab Gulf Programme for United Nations Development Organizations (AGFUND), the Arab Open University (AOU) started its actual operation with branches in six different Arab countries: Kuwait, Jordan, Lebanon, Bahrain, Egypt and Saudi Arabia. AGFUND officials signed an understanding agreement with the UK Open University (UK-OU) to benefit from its long-standing experience in the business of open learning. In addition to consultations, the agreement grants AOU a license to use UK-OU course materials subject to certain rules and standards to guarantee a high quality education.

In its first academic year in operation, AOU adopted the criterion-referenced approach or fixed (straight) scale to assign letter grades to students. On the basis of this grading method, grades typically are grouped into fixed intervals and a letter grade is assigned to correspond to each interval (See Appendix I). Based on a thorough investigation of the performance of students in first-year courses as reflected by their grades, that method was abolished. It was found that grades were skewed strikingly to the lower intervals. Such a result was alarming for negative implications in the future.

With more Cs and Ds, it was predicted that a large proportion of enrolled students would satisfy the graduation requirements according to the UK-OU's system but would not be able to maintain a 2 grade point average (GPA) which AOU regulations require for graduation. AOU has enforced this minimum GPA requirement to meet

the accreditation and validation requirements set by the ministries of higher education in the countries in which AOU operates. What complicates the matter is that students are not permitted to repeat core courses licensed from UK-OU for the purpose of improving their grade point average. These courses account for nearly 75 percent of the total number of credits stipulated as a minimum for graduation. At the same time, these courses carry heavy weights ranging between 8 and 16 credits.

AOU academic officials took the issue seriously, and a set of grading methods were contemplated to reach an alternative. After serious discussions, the Standard Deviation Method (SDM) was selected and all branches were formally instructed to implement this grading scheme. Adoption of SDM implied a shift from an absolute grading system to a relative (norm-referenced) one. Has the shift to SDM been beneficial in remedying the problem? At the same time, a group of students have complained that their grade point averages had been harmed by the scales generated from applying the method. Are such complaints based on a solid ground? Is it a sound argument that SDM results in unfairly penalizing capable students and unjustly benefiting poor students? In a broader sense, is the act of institutionalizing SDM defensible?

The purpose of this study is to provide objective answers to these inquiries. The paper is organized into five main sections. Section I succinctly discusses the standard deviation method. Section II is concerned with the AOU experience with that method. The methodology followed to achieve the purpose of this study is disclosed in Section III. The results of the data analysis are revealed in Section IV. A discussion of the results of the study and suggestions are presented in Section V.

I. STANDARD DEVIATION METHOD

Various grading methods have appeared in the literature of educational measurement and evaluation. These methods have been categorized into two major groups: relative grading methods and absolute grading systems (Terwillinger, 1989; Ebel and Frisbie, 1991; Frisbie and Waltman, 1992; Haladyna, 1999). Examples of the relative grading methods include the curve, the distribution gap, and the standard deviation. Fixed percent scale, total point, and content based are all examples of the absolute grading methods. Compared with other grading systems, the SDM is regarded as the soundest and the fairest in terms of assigning letter grades objectively though it is the most complicated computationally (Frisbie and Waltman, 1992). As a relative grading method, SDM is grounded primarily on the assumption that grading is needed for identifying students who perform best against their peers and to weed out the unworthy (CTL, 2001). According to this method, student grades are based on their distance from the mean score for the class. The distance is measured by using the standard deviation of the scores, which describes the variability or spread of scores around the mean score. In this sense, the standard deviation is employed for the sole purpose of determining cutoff points for each grade level.

To explain how SDM works and for a matter of simplification, let us assume that our grading system consists of five letter grades: A, B, C, D and F. To determine the cutoff points for these five letter grades, the mean and standard deviation for the raw grades need first to be calculated. Table 1 reveals an example of the equations needed to compute the lower scores of each interval corresponding to the specified letter grades. As shown in the table, cutoff points for "C" grades range from one-half standard deviation below the mean to one-half above, leaving the mean score to represent the midpoint of the C interval. The A-B cutoff point can be determined by adding one standard deviation to the upper C cutoff point. Subtracting one standard deviation from the lower C cutoff point would yield the D-F cutoff point.

Table 1
Example of Standard Deviation Method

Letter Grade	Lower Grade
A	$\mu + (1.5)\sigma$
B	$\mu + (.5)\sigma$
C	$\mu - (.5)\sigma$
D	$\mu - (1.5)\sigma$
F	*0*

For more clarification, let us assume that we have 200 students enrolled in course *x*. After a calculation of the mean and standard deviation of the scores of these students, we obtained $\mu= 65$ and $\sigma= 10$. Using the equations disclosed in Table 1, the resulting scale for the entire course is shown in Table 2.

Table 2
Scale Based on SDM for Course *x*

Letter Grade	Scale
A	80 – 100
B	70 - 79
C	60 – 69
D	50 – 59
F	0 - 49

This example exposes a straightforward application of SDM. However, in practice, ranges can be constructed smaller or larger for groups of higher or lower letter grades. This can be achieved by adjusting the number of standard deviations employed to determine the cutoff points. The grader may select the D-F cutoff to be one standard deviation below the mean and the A-B cutoff might be1.3 above the mean, rather than 1.5. For this, it should be noted that *SDM is not always performed in a mechanistic fashion.*

SDM involves the value judgment of the grader; and thus there is some arbitrariness in the method (Guskey, 1994; Gronlund, 2003). The value judgment of the grader plays a primary role in deciding whether the mean is set to C or B or whether the mean represents the lower or upper limit of the letter grade selected. The standard deviation distance between letter grades also is determined arbitrarily (Strashny, 2003). The grader may decide to have the lower limit of the B interval equals 0.3 standard deviation above the mean if he or she envisages the grades to be highly unreasonably concentrated in the C interval. The concentration of grades in a certain interval is usually measured by the percentage of students receiving the letter grade corresponding to that interval. Although SMD "requires no fixed percentages in advance, … the cutoff points are not tied to random error...[The grader must have] some notion of what the grade distribution should be like" so that some trail and error would be used to decide on how many standard deviations are needed to determine the cutoff points (Frisbie and Waltman, 1992).

II. AOU EXPERIENCE WITH SDM

During the examination period of the first semester of the academic year 2003-2004, all AOU branches were officially instructed to apply the SDM to the final grades prior to submitting them to the AOU Central Examination Committee (CEC) for approval. In exception of one or two branches, the instructions were not followed mainly due to the lack of experienced individuals who acquire the capability to perform the method properly. This academic year, with the insistence, continuous assistance and monitoring of CEC, all branches have succeeded in varying degrees to assign letter grades on the basis of SDM. This means that the straight or fixed scale in which grade intervals are predetermined in advance is no longer in effect.

The AOU grading system consists of seven letter grades. These grades and their numeric equivalents are displayed in Table 3. The third column of Table 3 represents percentage ranges corresponding to each letter grade. These percentages are not fixed; they are only used as guidelines to be taken into consideration in the process of distributing letter grades.

Table 3
AOU Grading System

Letter Grade	Points	Percent
A	4.0	10% ± 3
B+	3.5	30% ± 5
B	3.0	
C+	2.5	40% ± 5
C	2.0	
D	1.5	10% ± 2
F	0	

One distinctive aspect of the AOU grading system is that there are two types of the F grade. The first type is assigned to students who have passed the final examination, but their total score is below 50 or the minimum grade provided by SDM. The second type, denoted F*, is assigned to those who have failed to achieve 25 out of 50 in the final examination regardless of their total score. This is a requirement stated clearly in the AOU academic rules and regulations. The scores of students ascribed to the second type are excluded when SDM is applied. F grade is assigned to them automatically. When the percentage of those students receiving F* is quite high, let us say more than 20%, an investigation is launched to unveil what contributed to the occurrence of such a percentage. AOU administration relies on such information to take corrective actions and avoid similar experience in the future.

Maintaining a grade equal to or greater than 50% of the grade assigned to the final examination has created an obstacle for the straightforward application of SDM. According to AOU academic rules and regulations, a student *must* pass the final examination as a condition for passing a course. Thus, as a matter of fairness, a student who meets that specified condition and whose total score is equal to or greater than 50% of the total grade of the course *should* pass. It is unfair to accept to fail a student whose total is, say, 58 mainly because he achieved 22 on the final examination, and at the same time fail another student who received more than 50 of the total grade but less than the D-F cutoff point.

The implication of this argument is that AOU has tended to accept all grades lower than the D-F cutoff point up to the grade 50. However, if the D-F cutoff point goes lower than 50, that grade would mostly be accepted as a passing grade. To clarify the first case, let us assume that we obtained $\mu= 78$ and $\sigma= 11$ in a certain course. On the basis of SDM, the D-F cutoff point should be 61, meaning that any student who achieved 60 or less would receive F grade. This is true in a normal practice where there are no administrative constraints that demand an ethical judgment. As SDM practiced at AOU, grades below 61 up to 50 would be included in the D interval, resulting in pulling lower points of higher grades downwardly. That is, instead of having A to start from 95, it could start from 90, and that would affect the location of the lower grades of the rest of the intervals. In this sense, it could be stated that AOU does not employ a pure form of SDM, but a modified version of SDM that adequately harmonizes with its rules and regulations.

From experience, it has been noticed that AOU branches are not uniform when performing SDM. Since arbitrariness is widely accepted in the practice of SDM, such inconsistencies are expected. Different graders hold different notions of how grades should be distributed, logically producing dissimilar results. One major question pertaining to SDM is: should the mean be C or C+ or B? Another question is: should there be an equal number of D and B+ or A grades? Such inquiries are of a paramount importance to SDM application, but unfortunately can not be resolved on a scientific ground. That is, objectivity is overwhelmed by subjectivity. A review of the results submitted to CEC this semester reveals that one branch consistently treats the mean of the total scores as the lower point of the C+ grade while another assigns the mean as the lower point of the C grade. The results of other branches show that the mean falls somewhere within the C+ interval and mostly towards the upper limit of the interval. Some branches are concerned with maintaining closeness between the number of D and B+ grades while the concern of other branches is with balancing between the D and A grades.

III. METHOD

To achieve the purpose of this study, the final letter grades of the 2005-2006 academic year in Kuwait branch were used as raw data. Two main reasons lied behind the selection of data from Kuwait branch. First, Kuwait branch comprises more than one third of the total number of AOU student body. For this justification, Kuwait branch is assumed to be adequately representative of the total population of AOU students. The second justification is that Kuwait branch is geographically convenient; and thus this advantage facilitated full accessibility to the data. The data were limited to 18 courses licensed from UK-OU and were offered in that academic year.

The data set contained 10232 records which represent all students who attended the final examinations of the selected courses. Out of these records, 917 records were excluded since they comprised those who had not attended the final examinations with or without acceptable excuses plus those who had disappeared for their own reasons. The remaining valid records were 9315. For the purpose of data analysis, the records of those students who failed the final examinations (scored less than 25 out of 50) were eliminated since their grades were not subjected to SDM. In doing so, the number of used records in the analysis was 8105 which represents 79.2% of the total records initially selected. The analysis was conducted using SPSS software, mainly frequencies, descriptive statistics, crosstabulation, and the nonparametric sign test.

IV. RESULTS

Data analysis was based on treating the records as a one set instead of analyzing them per course. This macro approach is deemed more appropriate for serving the purpose of the study. Descriptive statistics of the selected courses are presented in Table 4. The mean scores ranged between 65.2 for C18 and 78.7 for C11. The calculated grand mean for all scores was 72.4 ($SD= 11.4$). The mean score for female students was 73.2 ($SD= 11.4$; $n = 4489$ (55.3%)); and for male students, the mean was 71.3 ($SD= 11.4$; $n = 3616$ (44.7%)). The difference between those two mean scores was statistically significant ($t= -7.3$, $p\leq .000$). Detailed descriptive statistics and tests of the differences in the performance of male and female students in individual courses are disclosed in Appendix II.

Table 4
Descriptive Statistics of Selected Courses

Course*	Mean	SD
C1	73.5	9.3
C2	65.3	10.1
C3	69.9	9.9
C4	74.4	12.9
C5	69.7	9.8
C6	76.1	10.9
C7	66.8	10.4
C8	72.9	10.1
C9	71.9	12.3
C10	75.6	10.1
C11	76.9	10.3
C12	78.7	12.6
C13	75.8	12.0
C14	71.8	11.8
C15	72.3	9.8
C16	72.1	11.9
C17	71.4	10.5
C18	65.2	10.9

* *All course codes are pseudo.*

The average percentage of success in the selected courses was 83.8%, which is reasonably acceptable. The number of students attended the final examination ranged between 28 for C8, which is a core course in the English program, to 702 for C17, which is a university required course. All standard deviation values were reasonable and did not highlight significant variability in the grades of single courses.

Table 5 reveals a comparison between the grade distribution resulted from SDM with the distribution of grades that we would have obtained had the old fixed scale been used. It is clear that both methods produced almost similar distributions in respect to the D and F letter grades. Applying SDM resulted in lowering the C and C+ letter grades by 7.1% and 6.3% respectively and increasing the A, B+ and B letter grades by a total of 13.4%. The most striking result is that the total number of those students who received the letter grade D increased negligibly, which contradicts inspirations and expectations. But it seems that no further assistance could have been offered to these students. They received what they really deserved taken into account that out of those who passed the final examinations 1112 students (13.7%) scored between 25 and 26 points (marginal cases).
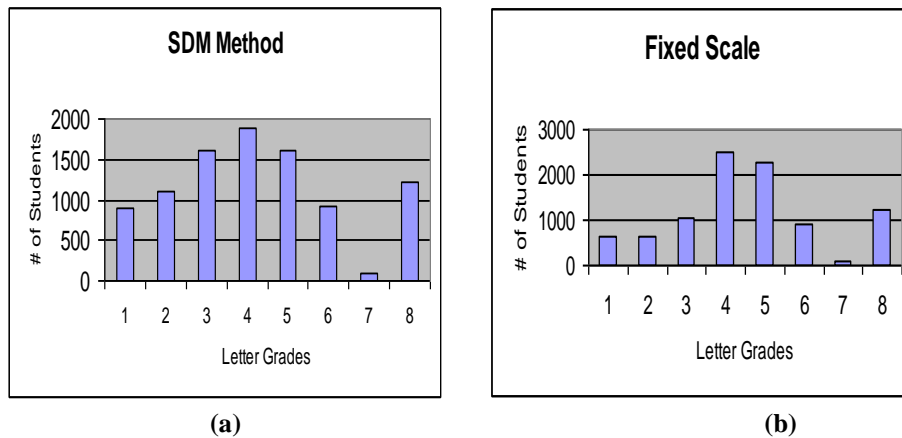
Table 5
Comparison between Grade Distributions of SDM and Fixed Scales

Letter Grade	Fixed Scale		SDM Scale	
	n	*%*	*n*	*%*
A	641	6.9	886	9.5
B+	641	6.9	1106	11.9
B	1067	11.5	1610	17.3
C+	2478	26.6	1890	20.3
C	2268	24.3	1606	17.2
D	916	9.8	925	9.9
F	94	1.0	82	0.9
F*	1210	13	1210	13

F* denotes students who failed the final examinations

Figure 1 presents two bar graphs of the data shown in Table 5. Figure 1.b illustrates that the grades obtained from the fixed scale are relatively skewed to the left (skewness is negative (-0.17)). However, the Kolmogorov-Smirnov test for normality was highly significant ($Z= 3.4$, $p \leq .000$), indicating a rejection of the null hypothesis that the distribution of the grades is not different from normal. Applying SDM to the grades created an adjustment that made the distribution close to normal as shown in Figure 1.a.

Figure 1
Grade Distributions of SDM and Fixed Scales

(a)　　　　　　　　　　　　　　　　(b)

Key: 1=A; 2=B+; 3=B; 4=C+; 5=C; 6=D; 7=F; and 8=F*.

A crosstabulation between the distribution of grades produced by both SDM and Fixed scales is revealed in Table 6. The diagonal cells represent the numbers of students whose grades did not change. That is, those students would receive the same letter grades regardless of which method is used. Cells fall to the right of each one of the diagonal cells include the numbers of those students whose grades were lowered as a result of applying SDM. On the other hand, cells fall to the left or below each of the diagonal cells designate the numbers of students who benefited from SDM. The most significant positive change occurred to students whose grades improved from C+ to B (*n* = 907) followed by improvements from C to C+ (*n* = 841). The highest negative change occurred to a group of students whose grades were lowered from C+ to C (*n*= 314).

Table 6
Crosstabulation of SDM and Fixed Scales' Grade Distributions

Fixed Scale	SDM Scale							Total
	A	**B+**	**B**	**C+**	**C**	**D**	**F**	
A	540	101	0	0	0	0	0	641
B+	246	303	92	0	0	0	0	641
B	75	419	474	99	0	0	0	1067
C+	25	283	907	949	314	0	0	2478
C	0	0	137	841	1050	240	0	2268
D	0	0	0	1	242	673	0	916
F	0	0	0	0	0	12	82	94
Total	886	1106	1610	1890	1606	925	82	8105

Note: *F* is not displayed.*

A nonparametric sign test was conducted to classify the changes presented in Table 6. To add more strength to the study, two separate sign tests were performed, one for male students and the other for females. The results of these three tests are revealed in Table 7. All three tests are statistically significant, confirming that differences in distributing grades exist for the two grading methods. As shown in Table 7, the grades of nearly 50% of students did not change as a result of applying SDM. Generally, SDM produced improvements in the grades of nearly 39.4% of all students who passed the final examinations of the selected courses (*n*=3188 students). More than 56% of the students benefited from SDM were females. At the same time, applying SDM resulted in lowering the letter grades of nearly 10.4% of students (*n*=846) more than half of whom were males (51.2%).

Table 7
Sign Test Results

	Male		Female		Total
Difference	**N**	**Row%**	**N**	**Row%**	**N**
Positive	1399	43.9	1789	56.1	3188
	(38.7)		*(39.8)*		*(39.4)*
None	1784	43.8	2287	56.2	4071
(Tie)	*(49.3)*		*(51.0)*		*(50.2)*
Negative	433	51.2	413	48.8	846
	(12.0)		*(9.2)*		*(10.4)*
Total	3616	44.6	4489	55.4	8105
Z	-23.7		-29.3		-36.9
p	.000		.000		.000

Note: Numbers in parentheses are column percentages

V. DISCUSSION AND SUGGESTIONS

SDM has been used for the last four semesters in AOU to assign letter grades to students. It should not be mistakenly understood that AOU policy makers institutionalized this method for the sake of pushing up students' grades. Had AOU had an unconditional freedom over the management of the offered courses, the problem stated in the introduction might not have been recognized as such. AOU's freedom in making decisions pertaining to the management of offered courses is constrained by the agreement that AOU had previously signed with the UK-OU.

Though limited to Kuwait branch, the results revealed in the previous section show that applying SDM created significant adjustment to the shape of the distribution of the students' grades. This adjustment resulted in improvements in the grades of nearly 39.4% of the students which considerably exceeds the percentage of those who suffered from a drop in the letter grades they received (10.4%). This result warrants discussion.

At the individual level, those who suffered may believe that SDM is unfair and unjust. Most students still hold in their own minds the fixed scale which is predominantly utilized in secondary schools. In this sense, the fixed scale embodies the students' frame of reference. Thus, students first assess the grade they achieve on the basis of their own fixed conceptualization of grade distribution. This cognitive assessment spawns expectations. When the letter grades assigned to them by SDM match or are higher than their expectations, acceptance mostly prevails. Otherwise, disappointment is nurtured and subsequently complaints arise. To avoid such circumstances, AOU is urged to enhance the awareness of students about SDM.

As human beings, students are self-interest seekers and thus act in accordance with that premise. It is a rare case that students complain when their letter grades are higher than their expected ones. So, the issue is that students should fathom that if there expectations were not met in one course, these expectations would most likely be met in another. Students' expectations *ought to* be based on perception of devoted effort and performance rather than being sprung from a nonexistent scale. An emphasis on that point should be part of the awareness effort.

For the AOU administration, SDM should be regarded as a defensible method. The essence of SDM does not violate any right of a student. When a student scores 90 in a course and receives B+, she has no right to claim that she should have received an A. The B+ or whatever letter grade is assigned to her reflects her performance in the course as compared with that of all other students enrolled in the course. Again, 90 or whatever score a student obtains in a course does not symbolize *a thing* in its own. A score derives its meaning once a student's performance is compared with that of her peers, that is, once SDM is performed.

Another argument for SDM is that SDM, as the results have shown, has produced the greatest amount of good (improvement in grades) for the greatest number (the students). This defense is utilitarian in tone (West, 2003). The rationale behind the institutionalization of SDM by the AOU higher management stemmed from a sense of responsibility to do good to the students. However, when grading is at issue, there will always be poor achievers. This implies that doing the good is a relative matter; that is, good *cannot* be done to all in all circumstances and at once. There will always be winners and losers, and the results have classified students into such groups. These groups will invariably be present regardless of which of the known grading methods is espoused.

The most crucial finding that one might be concerned with is that utilizing SDM did not minimize the number of students who received the D letter grade. With SDM, the number was 925, and without it the number could have been 916. Twelve out of those who received D could have been assigned F had the old scale been used. Moreover, as stated above, there will always be students whose performance compared with their peers is at the minimum acceptable level; thus they deservedly receive Ds. Since the D grade was problematic from start, it could be suggested that *the letter grade D should be withheld in courses licensed from UK-OU.* In doing so, students who do not attain a C letter grade are offered an opportunity to repeat 8- and 16-credit courses.

This suggestion does not go without criticism. Since the vast majority of licensed courses are offered over two consecutive semesters, repeating a course implies a delay on the student part to complete the requirements of the program in which he or she is enrolled. Though contradicts the philosophy of open learning, most students are

ambitious to graduate within the shortest period of time possible. So a delay may negatively affect the student's motivation to remain in AOU pursuing higher education.

Let us now flip the coin and bear in mind that our concern is with low achievers. By withholding the D letter grade, a student is offered a hope to graduate. Which is better for a student to finish all the requirements, but denied the degree due to a GPA that is lower than 2 or to stay for additional couple of years and graduation is guaranteed from an academic stand point? The answer is quite obvious. Moreover, if the D letter grade is kept in the AOU grading scale, the student has no option but to stay in AOU for extra couple of years to raise his or her GPA to the mandatory level. So a delay in graduation is inevitable. All that argument is based on the assumption that improving GPA is feasible. In fact, considering the weight of the licensed courses, the achievement of that objective seems to be surrounded with obscurity.

To add more support for a rule that would enforce the withholding of the D letter grade, let us first examine how the AOU scale would appear. The scale would comprise the following letter grades: A, B+, B, C+, C and F. Omission of the grade D will result in a redistribution of letter grades, benefiting a good number of genuinely dedicated students. In this sense, the essence of the rule exemplifies a reward for the good rather than a punishment for the poor. The benefits for the poor have been highlighted earlier, and equality as a value dictates that AOU should bestow a reward upon the good. In the existing grade scheme, the highest percentage of grades concentrates in the C+ interval. With the suggested system, almost an equal concentration would move forward to the B interval. Is not that advantageous for the greatest number?

The defense for withholding the D letter grade should not be understood as a call for compromising the academic integrity of AOU. Let us give a glimpse at Table 8 which shows the intervals of the C letter grade in the courses licensed from UK-OU and examined in this study. It is clear that in exception of a couple of courses, the lower points of the C letter grade were 60 and above. This is within an acceptable norm for the C letter grade. It is worth mentioning that out of those who received Ds in these courses, 362 students (39.1%) attained exactly 25 in the final examinations. It is highly likely that these students were marginal cases and they were dealt with accordingly. If that is true, was good done to them?

Table 8
Intervals of the C letter Grade

Course	Lower Grade
C1	62 - 67
C2	58 - 63
C3	57 – 62
C4	61 – 66
C5	62 – 66
C6	64 – 69
C7	63 – 68
C8	61 – 64
C9	59 – 64
C10	65 – 72
C11	67 – 71
C12	62 – 71
C13	63 – 68
C14	60 – 67
C15	60 – 66
C16	60 – 66
C17	61 – 66
C18	55 – 59

For the sake of preserving the academic integrity of AOU, a proposed scale is presented in Table 9. It should be kept in mind that this scale is tentative in nature; it is intended to provide guidance rather than being followed as it stands. Also it should be remembered that grading is not value-neutral. In most cases, the C-F cutoff point should be *close* to the number resulting from deducting half of the value of the standard deviation from the mean. However, in some cases, a grader may decide to subtract one standard deviation from the mean or more in order to arrive at a reasonable C-F cutoff. For instance, if the mean of the grades of a group of students in a specific course turned out to be 76 and the standard deviation is 12, the grader may choose to have 64 or a little bit lower as the cutoff point instead of 70. In other cases, when the mean is low, the grader may make a choice to have the mean as the C-F cutoff point and then work the scale up. As an example, if the mean is 57 and the standard deviation is 18, it would be more acceptable from an academic point of view to accept the mean as lower grade of the C interval. See Appendix III for the results produced by the proposed scale as compared to those of SDM in respect to C17.

Table 9
Proposed Grading Scale for AOU

Letter Grade	Lower Grade
A	$\mu + (1.5)\sigma$
B+	$\mu + (1)\sigma$
B	$\mu + (.5)\sigma$
C+	μ
C	$\mu - (.5)\sigma$
D	$\mu - (1.5)\sigma$
F	0

CONCLUSION

The results showed that in comparison with the fixed scale, SDM is beneficial for AOU students. The analysis presented in this study showed that SDM meets the utilitarian criterion of doing the greatest amount of *good* to the greatest number. However, SDM has not contributed significantly to alleviating the problem for which it was initiated. SDM would not be fruitful unless the D letter grade in the courses licensed from UK-OU is omitted from AOU grading system. It is better for a student to have an opportunity to repeat courses than to pass all courses but is denied a degree for failing to maintain the minimum GPA requirement. AOU officials are urged to give this issue immediate attention; otherwise, some students would not have a chance to realize the goals that they have been pursuing since they enrolled in AOU.

ABOUT THE AUTHORS

Dr. Aleisa is the Arab Open University Registrar and an Associate Professor in the Faculty of Business.

Dr. Alhemoud is an Associate Professor in the Faculty of Business and a Visiting Professor of Marketing & Management at the American University-Sharjah.

REFERENCES

Center for Teaching and Learning (2001). Teaching at Carolina: Evaluation issues. http://ctl.unc.edu/he2.html☺
Available online: Eble, R. and Frisbie, D. (1991). *Essentials of educational measurement*, 5th Ed. Englewood Cliffs, New Jersey: Prentice Hall.

Frisbie, D. and Waltman, K. (1992). An NCME instructional model on developing a personal grading plan. *Educational Measurement: Issues and Practice*, 11, 35-42.

Gronlund, N. (2003). *Assessment of student achievement.* Boston: Allyn & Bacon.

Guskey, T. (1994). Making the grade: What benefits students? *Educational Leadership*, 52, 14-20.

Haladyna, T. (1999). *A complete guide to student grading*. Boston: Allyn & Bacon.

Looney, M. (2003). Facilitate learning with a definitional grading system. *Measurement in Physical Education and Exercise Science*, 7, 269-175.

Munk, D. and Bursuck, W. (1998). Can grades be helpful and fair? *Educational Leadership*, 55: 44-47.

Ornstein, A. (1999). Grading practices and policies: An overview and some suggestions. *NASP Bulletin*, 78, 55-64.

Strashny, A. (2003). A method for assigning letter grades: Multi-curve grading. Available online: http://econwpa.wustl.edu:80/eps/em/papers/0305/0305001.pdf

Terwillinger, J. (1989). Classroom standard setting and grading practices. *Educational Measurement: Issues and Practices*, 8, 15-19.

West, H. (2003). *An introduction to Mill's utilitarian ethics.* Cambridge: Cambridge University Press.

APPENDIX I

AOU Old Fixed Scale

Letter Grade	Intervals
A	90 - 100
B+	85 - 89
B	80 - 84
C+	70 - 79
C	60 - 69
D	50 - 59
F	< 50

APPENDIX II

Descriptive Statistics by Gender and Difference Tests

Course	Mean	SD	Male			Female			*KS*	*T*	*U*
			Mean	SD	Rank	Mean	SD	Rank			
C1	73.5	9.3	71.2	7.9	18.4	73.9	9.6	22.0	NS	NS	
C2	65.3	10.1	64.5	10.8	318.7	66.0	9.5	347.1	NS	NS	
C3	69.9	9.9	68.8	9.8	385.4	70.7	9.9	435.9	S		S
C4	74.4	12.9	71.2	12.1	51.3	77.5	13.0	68.5	S		S
C5	69.7	9.8	68.0	8.6	69.0	70.7	10.1	72.5	NS	NS	
C6	76.1	10.9	71.9	7.9	34.2	77.3	11.4	49.5	S		S
C7	66.8	10.4	65.2	10.4	211.4	68.0	10.3	243.5	S		S
C8	72.9	10.1	69.6	10.5	11.3	74.2	10.0	13.7	NS	NS	
C9	71.9	12.3	70.3	11.0	101.1	73.6	13.1	118.0	S		S
C10	75.6	10.1	73.7	10.6	53.3	77.0	9.6	63.4	NS	NS	
C11	76.9	10.3	77.0	9.6	77.5	76.8	10.9	77.5	NS	NS	
C12	78.7	12.6	78.4	12.7	48.4	79.0	12.6	51.2	NS	NS	
C13	75.8	12.0	74.6	10.9	91.4	76.9	12.8	103.5	S		NS
C14	71.8	11.8	72.0	10.8	172.4	71.7	12.4	170.8	NS	NS	
C15	72.3	9.8	70.6	9.7	87.5	73.5	9.6	104.5	NS	S	
C16	72.1	11.9	72.2	11.8	116.0	72.1	11.9	113.3	NS	NS	
C17	71.4	10.5	72.1	10.7	283.7	70.8	10.3	263.5	NS	NS	
C18	65.2	10.9	65.6	12.0	42.4	65.1	10.7	42.5	NS	NS	

KS= Kolmogorov-Smirnov Test of normality. *When the value of this test is significant, it means that the distribution of the variable is different from normal distribution. Thus, a nonparametric test is preferable; otherwise, a parametric test should be used.*
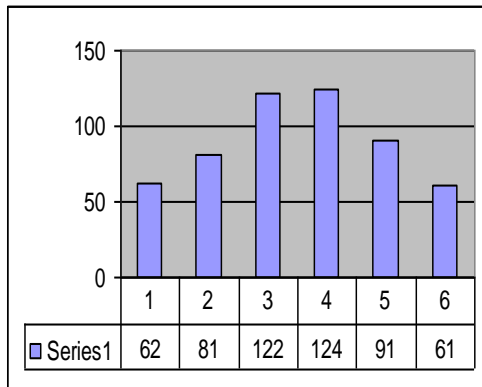*T*= T parametric test for differences between the means of two groups.
U= Mann-Whitney nonparametric test for differences between the mean ranks of two groups.
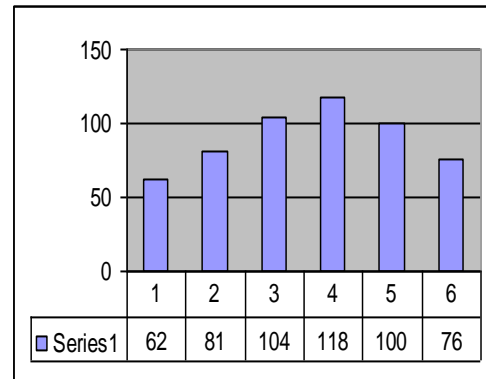**S**= Significant at *p≤.0*
NS= Not significant

APPENDIX III

Results of Proposed Scale as Applied to C17

	1	2	3	4	5	6
☐ Series1	62	81	122	124	91	61

Key: 1=A; 2=B+; 3=B; 4=C+; 5=C; and 6=F

	1	2	3	4	5	6
☐ Series1	62	81	104	118	100	76

Key: 1=A; 2=B+; 3=B; 4=C+; 5=C; and 6=D