Standardization of Exam Results in Premedical Years

A. El-Namrouty, F. Kheir, Sahar Algadi, Nawader Abdella

and Hiba Omer*

Department of Basic Sciences, Deanship of Preparatory Year, Imam Abdulrahman Bin Faisal University, Dammam, Kingdom of Saudi Arabia
** Corresponding author.*
E-mail address: aalnamrouty@iau.edu.sa, fkhair@iau.edu.sa, saaahmed@iau.edu.sa, neabdella@iau.edu.sa, hbomer@iau.edu.sa

ABSTRACT

The achievement of students in university's preparatory/ foundational years are influenced by the nature of different subjects they study as well as the difference in their background knowledge and skills acquired in secondary schools in both rural and urban cities. This has an impact on the competition among students who need to get high scores (B or above) in order to be accepted to the college of their choice. This is an important issue in many universities adopting preparatory/foundational year's system. This system is pretty much similar to pre-university colleges in UK and USA.

This work aims at providing ways to standardize the results and minimize the external elements affecting the students' scores. A few articles are found about standardization of exams rather than standardization of results in particular. Different standardization equations are applied to the results of the physics exam. The impact of these equations on A-students, mediocre students and weak students are analyzed.

Although standardization of results improved the scores of poor students, some methods reduced the scores of high students, which may affect their chances of competition especially if they are stronger in one of the main basic science subjects and weaker in others. Performing the same procedure on the results of the three subjects (Physics, Chemistry and Biology) will give us a better view of the situation and a chance to choose the most appropriate method of standardization.

Keywords: Standardization of exam results, Discrepancy in standards, ranks, mean, standard deviation

1. INTRODUCTION

Examinations are acceptable means to assess the cognitive skills and educational attainment of students at different levels. This allows schools to improve academic performance and might also improve cognitive skills (Finn ***et al.***, 2014). The scores in such exams may determine the future enrollment of the students into different colleges. Competition for such colleges may be based on the results of the students on their General Certificates Exams adopted in many countries or after completion of A-level exams or foundation or preparatory year. Students accepted in a preparatory/ foundational year are exposed to the same educational environment after which they compete to be accepted to the colleges of their interest. A preparatory year can help ensure a smooth transition into higher education from

secondary school. (AC, 2019) It reduces the discrepancy among schools in terms of infrastructure, facilities and quality of education.

An example of this case is the foundation or preparatory year program implemented in Imam Abdulrahman Bin Faisal University in Saudi Arabia which endorsed this strategy several years ago. It was originally created to promote collaboration and integration between three faculties; Architecture, Engineering and Design and continued for other tracks. (IAU, 2019) Students are accepted into disciplines or tracks including: Health, Engineering, Science and Humanitarian tracks. Health Track is the foundational year for the medical colleges: Medicine, Dentistry, Clinical Pharmacy, Applied medical Sciences and Nursing. The goal of the majority of students is to be accepted in the colleges of Medicine or Dentistry. Nevertheless acceptance to these colleges has strict requirements. Students should get a high collective score and not less than 80% in Biology, Chemistry and Physics. A weakness of comparison of raw scores in such subjects is that it does not adjust for weakness in the language, a deterministic element in Biology, or Mathematics, which is very influential in Physics. Scores obtained by students in a certain subject can indicate the proportion of the over-all mark the student gained. Nevertheless, these scores do not account for factors such as the difficulty of the test and how the student compares to his/her colleagues neither do they account for the margin of error in the test score. (NFER, 2017)

Many articles that discuss the standardization of exams and exam questions are available, but only a few investigated the standardization of results, although the issue was discussed in the early 40's by Quazi Hossain (Motahar, 1940). This is an important issue, especially for science subjects taught in a language other than the media of instruction in high schools. Standardization of results is also of vital importance where exams are marked manually, which may lack consistency in marking amongst faculty members (AlphaPlus, 2014)

The presented work studies the different standardization methods and a curve-up method and presents the outcome of the results with the different methods.

It aims at reaching a generalized standardization method accepted by colleges or universities implementing foundational years.

2. METHODS

An intensive analysis was performed using the results of a physics exam marked out of 15% using the standardization methods explained below. Further analysis for the normalized and standardized results was performed using fixed average: $\bar{X}$ and standard deviation: s for both male and female students later referred to as NX and STX. The results are displayed in tables; analysis was performed using the results of: A-students (14/15), mediocre students (7-12/15), and weak students (5/15). Students scoring below 5 out of 15 were excluded from this study since an increase in the results may unnecessarily raise scores of very poor students. For those, remedial sessions were conducted to improve their standards and comprehension of the subject.

Standardization Methods

Standardization of exam results, which can be described as: using the same yardstick for measuring otherwise different quantities presents a solution to such a dilemma. There are different standardization methods: these include:

1. **Using Ranks**:

The steps for standardization using ranks are summarized as: Rank the students in each subject/quiz separately and then in their totals. Add the ranks and compare these sets of ranks.

2. Using percentiles

This is the rank of a student per cent of the total number of students. It is calculated using the following equation

$$Percentile\ rank = \frac{Rank}{number\ of\ students} \times 100\% \quad (1)$$

3. Using Normalization

The word "normalization" is used informally in statistics, and so the term normalized data can have multiple meanings. In most cases, normalizing data eliminates the units of measurement for data, making it easier to compare data from different places. (Horse, 2019) The normalized method used here is transformation of data using a common z-score. Z-scores express scores not in raw points, but in relation to how everyone else on the test did. A standard or z-score measures the number of standard deviations between a raw score and the mean in standard deviations. This way standard score is defined as the distance between a raw score and its mean, in standard deviations. The z-score has a mean of zero; the marks should be adjusted by adding the mean result of students to the calculated z-value.

The standard or z-score is a very useful statistic tool because it (a) enables calculation of the probability of a score occurring within the normal distribution and (b) allows comparison of two scores that are from different normal distributions. (L. R 2019)

Factors needed to calculate the standard score include: the mean and the standard deviation of that distribution, as well as a raw score. Then we can apply the following equation:

$$z = \frac{X - X_{av.}}{s} \quad (2) \qquad \textit{(B, 2013)}$$

4. Using a fixed mean and standard deviations

This is another method for standardization of results, comparing two groups, for example boys and girls in the preparatory year, Health Discipline.

Standardization is performed using the following steps

a) Let X_1, X_2, …X_n be the raw grades of a class and Z_n the standardized score in the Physics examination. Let Y_1, Y_2,… Y_n be the raw grades and W_n the standardized score of a student in Chemistry of the same class – in the same order.

b) Calculate the two means: X_{av} and Y_{av} and the two standard deviations: σ_X and σ_Y

c) Standardize these results such that all examinations have the same mean: μ=50, for example and the same STD: s=15 (say).

$$\frac{Z_n - 50}{15} = \frac{X - \bar{X}}{\sigma_X} \quad (3)$$

So:

$$Z_n = 50 + \frac{15}{\sigma_X}(X - \bar{X}) \quad (4)$$

Similarly: $Y_n = 50 + \frac{15}{\sigma_Y}(Y - \bar{Y})$ (5)

5. **Using minimum competency level**

This system (which is not tested in this manuscript) was implemented at the Physics department, College of Medicine, University of Dammam. After the first draft of exam is prepared by the staff, each staff member examines the questions individually and puts a mark from 0-1 depending on the minimum competency level; this is an answer to: what is the mark expected to be obtained by the minimally competent student? An average of the marks stated by teachers is then obtained and the final result is multiplied by a factor decided from this calculation. There is a major drawback of this system, when applied in one subject rather than all. This will deprive the students who are stronger in this subject (who obtained 100%, which cannot be increased) from their chance to excel and thereby compete in the over-all assessment.

6. **Curve-up** This is mathematical edition of the result performed using the following equation

$$C = (10 \times \sqrt{X} - \frac{100 - X}{10}) \times 0.15 - 0.5$$

It is usually performed when the curve of the results is left-skewed to improve the results, such that the mean result of the students is above a certain value.

3. RESULTS

The rank of students scoring different marks (out of 15) is displayed in Table 1 below together with the percentile. Scores below 5.5/15 was not evaluated. Such students need special attention to improve in the subject. The total number of male students was 538 and that of female students was 454.

The different standardization methods are then applied to the raw data, as per score and the results are displayed in Table 2 below

Abbreviations:
R = raw data
C= curved-up results
N= Normalized
NX= Norm with fixed $\bar{X}$ & s
ST= Standardized
STX= Standardized with fixed $\bar{X}$ & s

Table 1: Score, rank and percentile

Grade	Males		Females	
	Rank	Percentile	Rank	Percentile
15.0	-	-	1	0.2
14.5	1	0.2	2	0.4
14.0	**7**	**1.3**	**5**	**1.1**
13.5	9	1.7	8	1.8
13.0	14	2.6	18	4.0
12.5	22	4.1	30	6.6
12.0	32	6.0	44	9.7
11.5	42	7.8	70	15.4
11.0	57	10.6	107	23.5
10.5	68	12.6	153	33.7
10.0	**86**	**16.0**	**186**	**40.9**
9.5	103	19.2	236	52.0
9.0	138	25.7	274	60.3
8.5	171	31.8	310	68.2
8.0	219	40.7	343	75.5
7.5	267	49.7	366	80.5
7.0	304	56.5	389	85.6
6.5	353	65.7	400	88.0
6.0	**383**	**71.2**	**419**	**92.2**
5.5	462	79.2	427	93.9

Table 2: Raw and standardized results as per score

Grade	Males No. of students					Females No. of students				
	R	C	N&ST	NX	STX	R	C	N&ST	NX	STX
15.0	0	0	6	0	0	1	0	0	1	1
14.5	6	4	2	6	6	3	1	1	3	3
14.0	2	2	5	2	2	4	6	3	4	4
13.5	5	13	8	13	5	10	22	4	22	10
13.0	8	10	20	10	8	12	14	10	14	12
12.5	10	25	15	10	10	14	63	12	26	14
12.0	10	11	11	15	10	26	46	14	37	26
11.5	15	35	16	11	15	37	83	26	46	37
11.0	11	35	17	35	11	46	37	38	83	46
10.5	18	33	35	35	18	33	36	46	38	33
10.0	17	96	33	33	17	50	56	33	36	50
9.5	35	37	48	47	35	38	23	50	33	38
9.0	33	49	48	48	81	36	11	38	23	69
8.5	48	30	86	86	48	33	19	0	34	23
8.0	48	43	30	30	37	23	8	36	19	23
7.5	37	34	50	43	49	23	7	33	8	11
7.0	49	20	34	34	30	11	2	23	7	19
6.5	30	22	20	20	43	19	2	23	3	8
6.0	43	16	22	38	34	8	4	11	6	7
5.5	34	8	16	8	20	7	2	19	2	2

4. DISCUSSION

At school level, statistical standardization or moderation is used to compare achievements in different schools in many countries including UK (NFER, 2017), USA (Randall, 2008) and (GWA, 2012) and (TISC, 2014) In Australia it is used to adjust school marks so that the moderated school marks more fairly reflect the relative standards of achievement of the two groups of students.

From the raw data displayed in Figure 1 below, the female students performed far better than the male students. The majority of female students scored above 8/15. The distribution of the score is within the normal distribution. The graphical representation of male students is left-skewed. This has been the case for many years. Female students are generally more hard-working and have higher sense of competition.

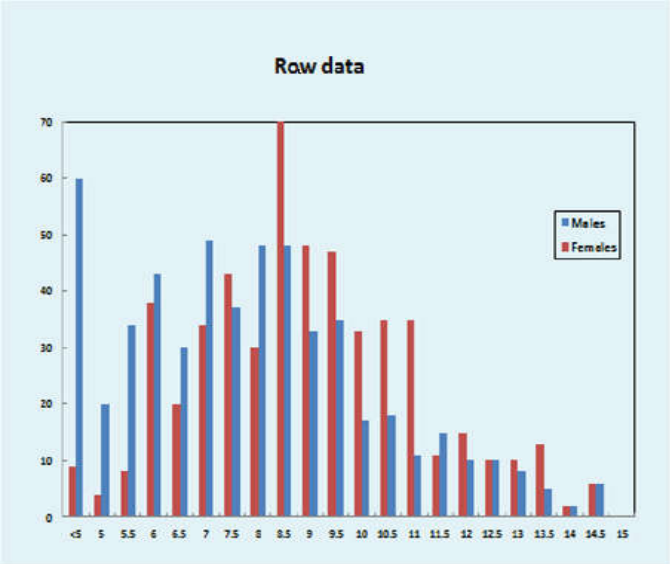

Fig. 1: Raw data for the results of male and female students

Figure 2 below shows the results after normalization using a fixed average and standard deviation. The scores have been increased tremendously. The distribution is now normal about 11/15 for female students and 8.5/15 for male students. This could raise an alarm that if implemented, some incompetent students may be accepted at colleges such as medicine and be responsible of the lives of their patients

Figure 3 below shows the effect of standardization using a fixed mean and standard deviation calculated from the results of the two categories.

Standardization using a fixed mean and standard deviation improved the results of the male students but reduced the results of females; since the average of female students was higher than that of male students.

This is in agreement with the findings of the online Tyari Team (OTT, 2018)

Multiplying the raw results by a certain factor, curve-up, generally improved the results of both categories except those who scored the full mark, as shown in Figure 4 below

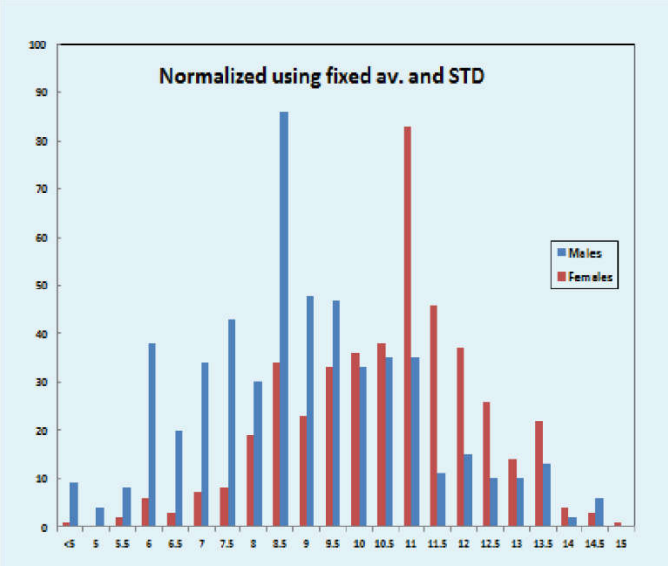

Fig. 2: normalized results

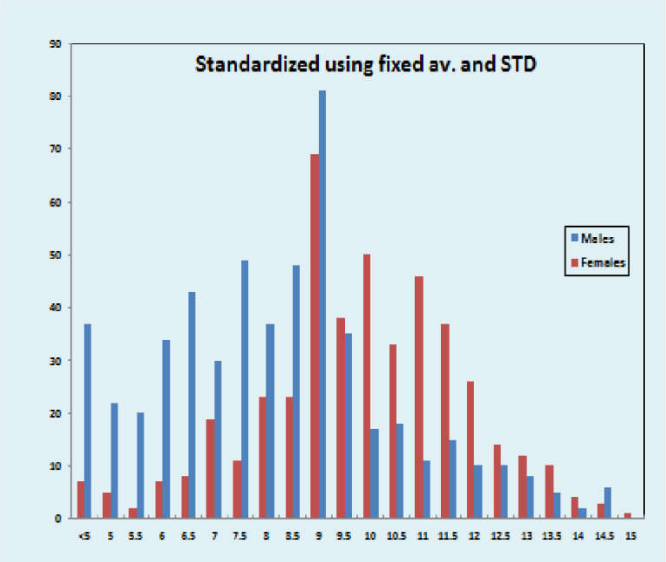

Fig. 3: standardized results using fixed mean and STD.

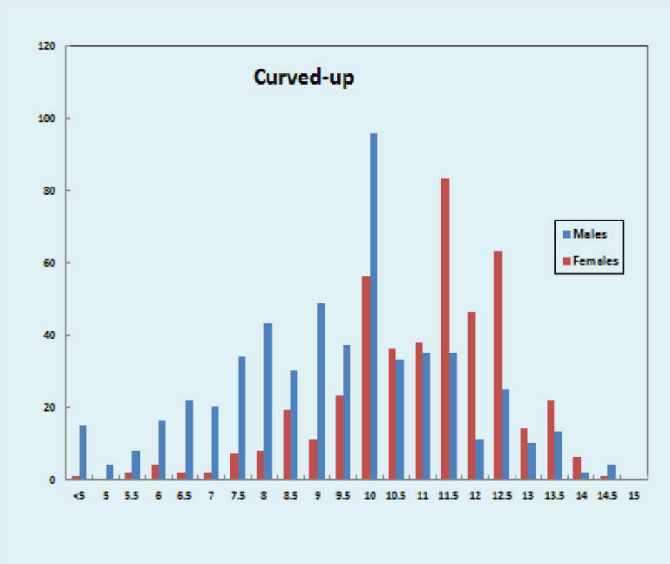

Fig. 4: curved-up results

Different equations applied on the raw data resulted in major changes in the scores. The effects depended on the row result. The effects on A-students, mediocre students and weak students were as follows:

A-students (14-15/15): Standardization using Ranks helps high standard students who performed poorly in one exam.

Normalization and standardization methods reduced the grades of "A" female-students. On the other hand, Standardization using a generalized average and STD does not affect the A-students. The curve-up reduced the results of the A+ female students who scored 15/15. N.B no male student scored the full mark.

Mediocre students (7-12/15): Standardization and curve-up may raise the results of mediocre students. Standardization using a generalized average and STD increases the grades and the number of students in this category

Weak students (5/15): Curve-up raised the grades, but not to the extent of passing the students.

The results can be summarized in the following table:

5. CONCLUSIONS

Standardization may be a useful tool in pre-medical college or any foundational year. It prevents the discrepancy in the standards resulting from the large number of groups taught by a variety of teachers; and the different levels of difficulties of the exams. However, this work shows that some standardization methods unnecessarily increased the results of poor students, and reduced the score of excellent students who had the full mark. This may affect the competition in case standardization was performed in one subject only especially if the student is stronger in one of the main basic science subjects and weaker in others. Performing the same procedure on the results of the three subjects: physics, chemistry and biology together, will give us a better view of the situation and a chance to choose the most appropriate method of standardization to implement in foundational pre-medical years.

Table 3: Effects of curve-up and different standardization techniques:

R	C	N(M)	N(F)	NX	ST(M)	ST(F)	STX(M)
15.0	14.5	-	14.5	15.0	-	14.5	15.0
14.5	14.0	15.0	14.5	14.5	15.0	14.0	14.5
14.0	**14.0**	**14.5**	**14.0**	**14.0**	**14.5**	**13.5**	**14.0**
13.5	13.5	14.5	13.5	13.5	14.5	13.0	13.5
13.0	13.5	14.0	13.0	13.5	14.0	12.5	13.0
12.5	13.0	13.5	12.5	13.0	13.5	12.0	12.5
12.0	12.5	13.0	12.0	12.5	13.0	11.5	12.0
11.5	12.5	12.5	11.5	12.0	12.5	11.0	11.5
11.0	12.0	12.5	11.0	11.5	12.5	10.5	11.0
10.5	11.5	12.0	10.5	11.0	12.0	10.0	10.5
10.0	**11.0**	**11.5**	**10.0**	**11.0**	**11.5**	**9.5**	**10.0**
9.5	11.0	11.0	9.5	10.5	11.0	9.0	9.5
9.0	10.5	10.5	9.0	10.0	10.5	8.0	9.0
8.5	10.0	10.5	8.5	9.5	10.5	7.5	9.0
8.0	10.0	10.0	8.0	9.0	10.0	7.0	8.5
7.5	9.5	9.5	7.5	8.5	9.5	6.5	8.0
7.0	9.0	9.0	7.5	8.5	9.0	6.0	7.5
6.5	8.5	9.0	7.0	8.0	9.0	5.5	7.0
6.0	**8.0**	**8.5**	**6.5**	**7.5**	**8.5**	**5.0**	**6.5**
5.5	7.5	8.0	6.0	6.5	8.0	4.5	6.0
5.0	7.0	7.5	5.5	6.0	7.5	4.0	5.5

REFERENCES

AC, A. C. (2019) *Compare 89 Preparatory Year Programs 2019* [Online]. Available: https://www.academiccourses.com/Preparatory-Year/Programs/ [Accessed].

ALPHAPLUS. (2014) *Standardisation methods, mark schemes, and their impact on marking reliability* [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/605660/2014-02-14-standardisation-methods-mark-schemes-and-their-impact-on-marking-reliability.pdf [Accessed 14 March 2019].

B, B. (2013) *Fundamental Statistics for the Behavioral Sciences* [Online]. Available: http://www.bryanburnham.net/ [Accessed 15 March 2019].

FINN, A. S., KRAFT, M. A., WEST, M. R., LEONARD, J. A., BISH, C. E., MARTIN, R. E., SHERIDAN, M. A., GABRIELI, C. F. O. & GABRIELI, J. D. E. (2014) Cognitive Skills, Student Achievement Tests, and Schools. Psych. Sc., 25**,** 736-744.

GWA, G.O.W.A. (2012) School Curriculum and Standards Authority, Statistical moderation of school marks.

HORSE, T. 2019. *Statistics How To : Statistics for the rest of us!* [Online]. Available: https://www.statisticshowto.datasciencecentral.com/normalized/ [Accessed 2 April].

IAU, I. A.B.F.U. (2019) *Deanship of Preparatory Year and Supporting Studies* [Online]. Available: https://www.iau.edu.sa/en/administration/deanships/deanship-of-preparatory-year-and-supporting-studies [Accessed 12 January 2019].

Lund Research (2019) How to do Normal Distribution Calculation [Online] Available https://statistics.laerd.com/statistical-guides/normal-distribution-calculations.php/ [Accessed 25 April]

MOTAHAR, H. Q. (1940) Standardisation of Examination Marks. *The Indian Journal of Statistics,* 5**,** 295-300.

NFER, T. N. F. F. E. R. I. E. A. W. (2017). The National Foundation for Educational Research in England and Wales Standardised scores and percentile rank
OTT, O.T.T. (2018) Do You Know The Effects Of Normalisation Of Marks in Government Exams? .
RANDALL, R. 2008. Teaching to the rating: School accountability and the distribution of student achievement J.of Pub. Eco., 92.
TISC, T. I. S. C. 2014. Marks Ajustment Process for University Admission.