TECHNICAL ISSUES IN GRADE-SCORE CONVERSIONS FOR A TRANSITION FROM NORM REFERENCING TO CRITERION REFERENCING

Wing Hung Wong[1], Kam Moon Pang[1]*, Shelly Liang Liao[2]

[1] Office of University General Education, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China

[2] Baldwin Cheng Research Centre for General Education, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China

Abstract

Criterion referencing evaluates a student's performance in an assessment against a predetermined set of goals and standards, while norm referencing measures a student's performance by comparing it with the performance of other students. Although many universities in Asia-Pacific have already transited from the historically adopted norm referencing to criterion referencing, some technical issues have yet to be fully addressed. In this paper, we shall argue that the adoption of criterion referencing does not imply a dismissal of normal distribution and hence not norm referencing; and demonstrate an application of mathematical justifications to legitimating a grade-score conversion so as to avoid a rampant grade inflation or deflation during the transition from norm referencing to criterion referencing. This, in addition, leads to a reflection on the legitimacy of the linearity of Grade Point Average.

Keywords: *criterion referencing; norm referencing; grade-score conversion; normal distribution; grade-score equivalence*

1. INTRODUCTION

Assessment plays a critical role in university education and serves a twofold purpose. Students can understand their strengths and weaknesses, identify areas for improvement, and thus be motivated to learn and grow. On the other hand, the university can evaluate the predetermined learning outcomes, measure teaching effectiveness, and thus ensure the quality of education.

Norm referencing has been historically adopted in schools and colleges, as well as in public examinations. Norm-referenced assessments compare an individual's performance to the performance of a larger group of students who have taken the same test; and provide percentile scores that indicate where the individual falls in relation to the group, allowing for comparisons and ranking. The grades awarded to students are transformed from the percentile values and are subject to 'curving' to fit a prescribed grade distribution. Hence, norm referencing is a useful tool to identify the students who perform the best relative to their peers. It is commonly used for standardised tests, such as the SAT, ACT, and other professional certification exams; and also for ranking and selecting students for specific programs or opportunities such as scholarships. Yet, such assessment offers little information about what individual students have actually learnt and how such learning has taken place because it is a relative grading. Thus, norm referencing does not achieve all of the purposes of assessment in university education on its own. Furthermore, the transformation from the percentile values to grades during curving is arbitrarily chosen with no legitimacy provided. Sadler commented that this grading approach "[had] long been criticised as logically flawed in principle and relying on untested statistical assumptions" (Sadler 2009).

In contrast, criterion referencing can achieve the full purposes. Criterion-referenced assessment measures student performance against a set of predetermined goals and standards, which are collectively referred to as 'criteria' hereafter. These criteria are the sole consideration in grading. The grades thus awarded are not subject to 'curving' or adjustments to fit a prescribed grade distribution. Ideally, the criteria are made known to students well in advance to offer a good sense of purpose and self-understanding —what knowledge, skills and attitudes the student may aim to achieve (goals intended);

where the student is standing in achieving them (standards met); and how the student may move up the "ladder" of achievements. Since criterion-referenced assessment has an obvious advantage that it could provide both the teacher and student with information about the strengths and weaknesses of student performance, more colleges and universities have considered or accomplished making a transition from norm referencing to criterion referencing. It can also be observed that more educational institutions and examination bureaux have taken criterion-referenced assessments into consideration when evaluating student competences (Carlson et al. 2000; Bloxham, Boyd & Orr 2011).

Transiting from norm referencing to criterion referencing calls for alignment of assessments with learning outcomes; and development of appropriate criteria and in turn grading rubrics. In addition, revisions of grading standards and assessment methods are inevitable so as to ensure that they are aligned with some specific criteria. Criterion referencing in General Education Foundation Programme at The Chinese University of Hong Kong was launched 2018. The process of transition from norm referencing to criterion referencing is definitely a challenging task and has been extensively discussed (Cheung et al. 2018; Liao, Wong & Pang 2021, pp.151-166).

Some technical issues such as the legitimacy of grade-score conversions for the transition are equally important but always underemphasised. In practice, a course has different assessment components, e.g., quizzes, essays and presentations, and each of which carries a different weight, and a course grade is calculated by the weighted mean. Yet, a grading rubric, which is a grid of predetermined criteria, is devised in a criterion-referenced assessment. A student's grade would be determined by matching his/her performance against the criteria. However, criterion referencing itself does not offer any hint for the teacher to combine the grades of various components, let alone produce a course grade.

As the addition is not straightforward, one is tempted to adopt an augmented version of norm referencing as an intermediate step of the transition: 'curving' is still used for calculating the course grades as norm referencing does; in addition, each student receives a set of detailed feedback from teachers on his/her performance in every assessment component. However, there are two drawbacks: (i) The course grade still only reflects his performance relative to his fellow classmates. The matching between the course grade and his attainment of the predetermined criteria of the course is still a mystery. (ii) Without a grid which lists out the grades and the criteria, the student cannot see a ladder of achievement, let alone climb it. In other words, a full launch of criterion referencing is not complete unless a way to 'add' together the qualitative descriptions provided by criterion referencing is found.

In this paper, we shall focus on two technical issues arising in a transition from norm referencing to criterion referencing, with special attention paid to the logical connection with the context and mathematical properties of statistics. The discussion is outlined below. Firstly, the nature of the bell shape and its relation to criterion referencing and norm referencing will be discussed. Secondly, we shall consider a course with multiple sections and prove mathematically that similarly competent students would end up getting different course grades if the course teachers have adopted different grade-score conversion tables. Thirdly, we shall propose a legitimate grade-score conversion table by considering the mathematical properties of normal distributions. The paper will end with a conclusion.

2. JUSTIFIABILITY OF A SCORE DISTRIBUTION BY THE BELL SHAPE

Curved grading (in short, curving) is usually adopted in norm-referenced assessments. This method uses numeric scores to represent student performance, and the scores are arranged in order. Then, these scores are converted to percentiles, and the percentile values are transformed to grades according to a division of the percentile scale into intervals, where the interval width of each grade is in accord with a prescribed grade distribution. One may therefore regard that the distribution of the score values is less relevant in norm referencing. Yet, we shall demonstrate that the score distribution in a norm-referenced assessment reflects whether the assessment concerned is justifiable.

It has been noted that the distribution of student scores in an examination such as the SAT approximately follow a "bell shape". (Jaynes 2003, pp. 592-593; Dorans 2002) The "bell shape" is an informal name of the normal distribution. It represents a phenomenon in a quantitative manner. A normal distribution can be represented by a symmetrical curve over a data axis. It is in fact a histogram of a huge number of narrow bars. Figure 1 shows a typical example of a normal distribution in which *x*-axis and *y*-axis are, respectively, the scores in an assessment and the number of students. It shows, for every score, the number of students who got that score in an examination. Most scores in a normal distribution lie close to the central peak, which signifies the mean and the mode of the scores. The tails on the far sides represent the tiny portion of data which are far from the mean score. Moreover, it is reasonable to assume that the scores are a manifestation of certain student competences, then the horizontal axis of Figure 1 indeed represents the student competences.

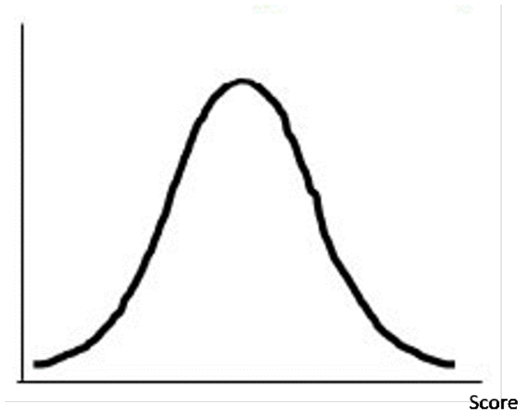

Fig. 1. A plot of a normal distribution

In cases where the score distribution deviates from the standard "bell shape", it might reveal that the assessment could not effectively assess and differentiate the student competences in concerned. Figure 2 shows a shifted distribution due to variations in the design of assessment, the so called right-skewed distribution. It is probably due to the question design in the assessment being so easy that most students are able to answer them correctly and thus will get rather high marks. Since the score distribution (dashed line) and the student competences (solid line) are quite out of alignment, a conversion of these scores to percentiles does not truly reflect the student competences.

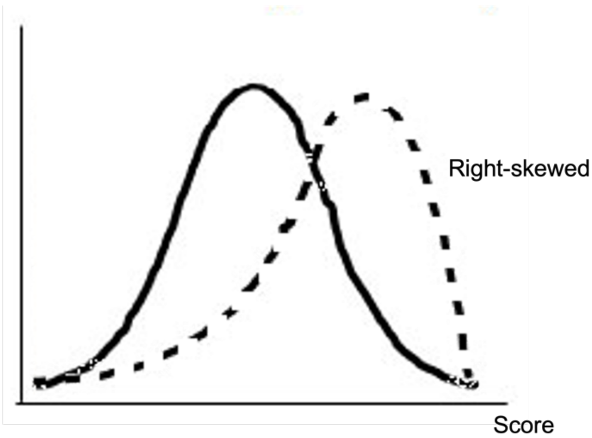

Fig. 2. The right-shewed distribution (dashed line) and students' competence (solid line)

In another extreme, if the question design in the assessment is so difficult that most students are not able to answer correctly, the score distribution will then be left-skewed. Another scenario is that if the levels of difficulty of the assessment are coarse, the distribution will become narrow because many students will fall into more or less the same category. In any of the above, the assessment mismatches the student competences in concern. Hence, if the grades are awarded to students mechanically according to a prescribed grade distribution and regardless of the shape of the distribution, the grades awarded would not fully and sufficiently differentiate the student competences. In other words, there is no statistical legitimacy in curving when one transforms the prescribed percentile values to grades regardless of the shape of a score distribution. Sadler's criticism on norm referencing is thus echoed and elaborated (2009).

3. MYTH: NO MORE BELL SHAPE IN CRITERION REFERENCING

Criterion referencing should also start with normal distribution. The most crucial step is to set suitable criteria that help students know their strengths and weaknesses, thus providing a scaffold for them to move up to higher levels of sophistication. The word "suitable" means that the predetermined criteria can match student competences. If the predetermined criteria are too lenient, less competent students could achieve high levels of sophistication. An illustrative example is that in a university level mathematics course, if one criterion is "the ability to do addition accurately", since most university students are able to do addition without difficulties, the criterion could not differentiate students' competence. On the other extreme, if the criteria for a doctoral thesis are used in a high school project, even a highly competent student might not be able to achieve the lowest levels of sophistication.

If no significant change in student competences in recent years is assumed and a transition from norm referencing to criterion referencing should not result in a significant change in students' grades, the performance of previous cohorts as assessed by norm referencing would be a good reference when setting suitable criteria. It is therefore in no way to say that the adoption of criterion referencing implies a dismissal of normal distributions or norm referencing (Lok et al. 2016).

When a student's performance meets the criteria of a certain level of sophistication, the corresponding grade and an accompanying score (e.g., 80-100 associated with A grade range) will be awarded. Since the criteria are all explicitly written down, the student would know how well he actually performed. In addition, the scores of all assessment components could be added up to a weighted sum, which would become a final grade of the course.

In short, as long as the criteria are well written by considering students' performance in previous cohorts, and the grade-score conversion is well justified by the bell shape distribution, the grade a student receives from criterion referencing should be commensurate with that from norm referencing. A smooth transition from norm referencing to criterion referencing can be achieved, and a grade inflation or deflation in the transition will not happen.

The building of criteria depends on individual course or programme. In contrast, we shall focus on technical issues of the legitimacy of grade-score conversion in which the underlying principles can be applied to any other teaching arrangements.

4. THE NECESSITY TO UNIFY GRADE-SCORE CONVERSION ACROSS SESSIONS

It has been noticed that universities might adopt different grade classifications. Some universities divide four levels of achievement in a course. A student is given a grade A for excellent performance, B for good performance, C for fair performance, and D for fail. The others prefer five levels such as grade A for excellent, B for very good, C for good, D for fair, and F for fail. Moreover, a course teacher might set a minimum score for grade A is 80, and the others set 95 is the minimum score.

Without loss of generosity, suppose a set of criteria are divided into 4 levels of sophistication, labelled (from highest to lowest) as A, B, C and F and the corresponding score ranges under a normal distribution are, respectively, 80-100, 68-79, 50-67 and 0-49. We reiterate that the number of levels of sophistication can be a convention, but the corresponding grade-score conversion is not. The justification of the preset score ranges will be explained in Section 5.

There are usually more than one assessment component in a course. As mentioned above, a common practice is to convert the individual grades of assessments into numeric scores, whose weighted sum will be finally converted back to a course grade. For a course with multiple sections where students are divided into different classes each taught by a different teacher who follows the same syllabus, a question arises immediately. Can each individual teacher adopt his own grade-score conversion table? The answer is "No".

Table 1 shows an illustrative scenario. Suppose there are two assessment components, each contributing to 50% of the final grade. Three students X, Y and Z, from sections taught by Teachers P, Q and R, respectively, and for an illustrative propose all get mid-A and mid-B in the two assessment components. Their performance, according to the same grading rubric, is considered the same. However, their teachers apply different grade-score conversion table and use different score ranges for the grades as shown in Table.

	Teacher P **A: 85-100** **B: 60-84** **(Student X)**	**Teacher Q** **A: 5, 5.5 or 6** **B: 3.5, 4 or 4.5** **(Student Y)**	**Teacher R** **A: 80-100** **B: 69-79** **(Student Z)**
Assignment 1 (mid-A)	92.5	5.5	90
Assignment 2 (mid-B)	72	4	74
Total score	82.25	4.75	82
Course grade	B	A or B	A

Table 1. Three students X, Y and Z, from sections taught by Teachers P, Q and R.

The three students of similar competence will end up getting different final grades due only to their teachers adopting different sets of score ranges. An even greater consideration for unifying the score ranges is that a change in a grade-range cut-offs is equivalent to a change in the assessment scheme. In other words, the weighting of each assessment component is changed (see the mathematical proof in Appendix). Hence, for a course with multiple sections, teachers of different sections must adopt the same grade-score conversion table. Otherwise, students taking the same course and getting the same grades in all components might end up getting different final grades. Students are likely to feel unfairly treated if they are aware of the situation. The discussion above could be extended to courses with more assessment components.

5. SUGGESTED GRADE-SCORE CONVERSIONS

In the previous section, we arrived at a conclusion that students with similar competences according to the same grading rubrics will undesirably end up getting different course grades, if teachers adopt different grade-score conversion tables. Hence, the same grade-score conversion must be adopted. However, there are multifarious choices of such a grade-score conversion. Among all possibilities, are there better ones? Having considered the mathematical properties of normal distributions, we suggest a conversion below, which is suitable for both single- and multiple-sectioned courses with multiple assessment components.

In the first place, the situation we are dealing with is a transition from norm referencing to criterion referencing. A student's grade should not significantly change due to the transition because only the description of the student's competence changes but not his competence. In other words, the assumption of a norm distribution of student competences in concerned would be the point to start with. Mathematical properties of normal distributions should thus be studied.

Although the actual grade distribution in a criterion-referenced assessment is in no way predetermined, a rampant grade inflation or deflation during the transition is not expected as long as the criteria are appropriately set. For simplicity, we consider a simple grading policy that students taking a course will receive either a grade A, B or C and that there is a consensus that the grade distribution of a course in question should not deviate drastically from Table 2.

Grades	Percentage of students	Ranges of *Z*-score
A	Top 30%	0.5205 – 3.0000
B	Next 50%	-0.8465 – 0.5205
C	Bottom 20%	< -0.8465

Table 2. A consensus of grade distribution and the corresponding *Z*-score ranges

Note that those students who fail in the course are excluded from our consideration because their performance might be incomparable due to reasons, e.g., student's medical problems and family matters other than their competences concerned in the course. Secondly, the interval between the passing score and the full score of the assessment is divided into 6 standard deviations since the interval covers more than 99.7% of students, and hence practically all students have been covered. By looking up a table of standard normal distribution, ranges of *Z*-scores for various grades can be easily calculated. In this case, the *z*-score cut-offs are -0.8465 and 0.5205.

Table 3 depicts two selected score-range scenarios for the consensus of grade distribution, i.e., 30% for grade A, 50% for grade B, and 20% for grade C: (i) A passing score range 50-100 is adopted. The standard deviation is $(100 - 50)/6 = 25/3$. The score cut-offs are $75 - 0.8465 \times 25/3$ and $75 + 0.5205 \times 25/3$, that is, 67.95 and 79.33, which are rounded upward to 68 and 80; and (ii) A passing score range 1.0-4.0 is adopted. The standard deviation is $(4 – 1)/6 = 0.5$. The score cut-offs are $2.5 - 0.8465 \times 0.5$ and $2.5 + 0.5205 \times 0.5$, that is, 2.08 and 2.76, which are rounded upward to 2.1 and 2.8, respectively. By following the same idea, the score cut-offs for any score-ranges and any grade distribution scenarios can be calculated.

Grades	Score range: 50-100	Score range: 1.0-4.0
A	80 – 100	2.8 – 4.0
B	68 – 79	2.1 – 2.7
C	50 – 67	1.0 – 2.0

Table 3. The score ranges corresponding to grade A, B, and C

In practice, there are always more than one assessment component in a course. As long as grade-score conversion tables have been properly constructed, "adding up" the assessment components is then mechanical. Here is an illustrative example. Suppose there are three assessment components in a course, with weighting 20%, 30% and 50%. If a student gets grades low-A, mid-B and mid-A in the three components. For each component, the teacher would look up a score in the score range that best reflects the student's performance. In scenario (i) in Table 3, the passing score and the full score are again 50 and 100, respectively, and the three scores of the three components are then 80, 75 and 90, the total score would be 80 × 20% + 75 × 30% + 90 × 50% = 83.5, which is in the score range of grade A (from 80 to 100). The student is therefore awarded grade A. In scenario (ii) in Table 3, the passing score and full score are, respectively, 1.0 and 4.0, then the scores of the three components are 2.8, 2.4 and 3.4. Therefore, the total score would be 2.8 × 20% + 2.5 × 30% + 3.4 × 50% = 2.98, which falls into the A range (from 2.8 to 4.0). In other words, the student will be awarded the same grade regardless of the score ranges being adopted, and thus the unfairness that was mentioned in Section 4 will not happen.

In addition, several points are worth noticing. The same principle is applicable to any student competences as long as they are described by a normal distribution; also applicable to any other score-range scenarios by simply adjusting the Z-score cut-offs according to the standard normal distribution table, and then the score cut-offs are recalculated accordingly.

Secondly, the number of standard deviations between the passing and the full scores would arrive at similar score ranges corresponding to grades. For example, five standard deviations are chosen, and a score range 50-100 is adopted, and the new standard deviation is in turn (100 - 50)/5 = 10, and then the score cut-offs are 75 - 0.8465 × 10 and 75 + 0.5205 × 10, that is, 66.6 and 80.2, which are rounded upward to 67 and 81, respectively. Therefore, the score cut-offs are very similar regardless of the chosen number of standard deviations as long as all students are essentially included. From a normal distribution table, it is found that the five and six standard deviations cover, respectively, 98.8% and 99.7% of students.

Thirdly, the score ranges corresponding to different grades are not necessarily equal in width. Levels of sophistication do not necessarily represent competences on a linear scale. This inspires us to reflect on the Grade Point Average (GPA) system which is commonly adopted in universities in Asia-Pacific. Letter grades A, B, C and D are equivalent to, respectively, GPA 4.0, 3.0, 2.0 and 1.0. However, such a linear scale of GPA can hardly be commensurable to usual practices of grade assignment. The incommensurability can be shown by *reductio ad absurdum*. In Figure 3, we plot the normal distribution of students' competences against the GPA. A standard normal distribution table would show that 2.4%, 30%, 60% of students are supposed to get A grade range, B range and C range respectively (Table 4). Apparently, this grade distribution is not commensurable to most courses or programmes partially because of gradual grade inflation in most universities in recent years. For example, the grade distribution of most courses in The Chinese University of Hong Kong is 30% for A grade range and 50% for B grade range. Thus, the incorporation of the linearity of the GPA in grade-score conversion calls for further justification.

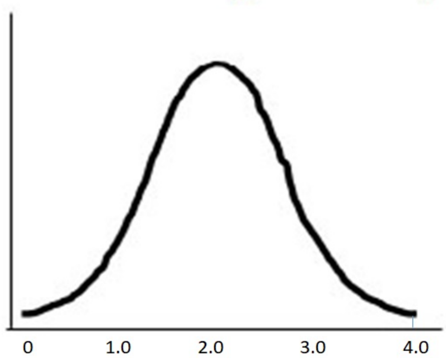

Fig. 3. A plot of students' competences against Grade Point Average system.

Grades	GPA	Percentage of students (%)
A- / A	3.7 – 4.0	2.43
B- / B / B+	2.7 – 3.3	30.08
C- / C / C+	1.7 – 2.3	60.68
D	1.0	6.68

Table 4. The grade distribution for the GPA, ranging from 0 – 4.0 and six standard deviations in total.

6. CONCLUSIONS

In contrary to norm referencing which distinguishes differences in student performance among fellow classmates, criterion referencing calls for the articulation of what teachers expect students to learn and the gathering of evidence to determine what students have attained and the learning ladder which students climb on. Thus, criterion referencing can achieve the full purposes of assessment in university education. Many universities especially in Asia-Pacific have already transited from the traditionally adopted norm referencing to criterion referencing, the authors argue that technical issues in grade-score conversions for the transition were yet to be fully addressed. Firstly, the adoption of criterion referencing does not imply a complete dismissal of norm referencing, because the predetermined criteria in a criterion-referenced assessment should match with students' competences which are related to normal distributions and in turn norm referencing. Secondly, "adding up" different components to generate a course grade in criterion-referenced assessments calls for a grade-score conversion, which are in fact not just a convention. The incompatibility of various grade-score conversions refers to a change in the weights of assessment components concerned. Thirdly, as long as a rampant grade inflation or deflation in the transition from norm referencing to criterion referencing is not expected, that means, an achievement of the consensus of grade distribution, the cut-off scores should be determined by the mathematical properties of normal distributions. It, in addition, it leads to a reflection on the legitimacy of Grade Point Average (GPA) because of the fact that most students get grades B or C in a course or a programme. Lastly, transiting from norm referencing to criterion referencing is doubtless a challenging task. Recently, generative Artificial Intelligence (AI) has had a significant advancement, and it might help solve the challenges in the transition and in the implementation of criterion referencing. For example, AI can generate assessments that align with specific learning outcomes, ensuring that the assessments are valid and reliable measures of students' performance in the assessments concerned. AI might also help score assessments accurately according to a pre-determined grading rubrics, providing rapid feedback to students (Baidoo-Anu & Owusu-Ansah 2023). This can be a booster of adopting criterion referencing.

ACKNOWLEDGMENTS

The authors would express their gratitude to Professor Kenneth Young for fruitful discussions and to Dr Julie Chiu for inspiring us to write on this topic.

REFERENCES

1. Bloxham, S, Boyd, P & Orr, S 2011. "Mark My Words: The Role of Assessment Criteria in UK Higher Education Grading Practices." *Studies in Higher Education*, vol. 36, iss. 6, pp. 655-670.
2. Carlson, T, MacDonald, D, Gorely, T, Hanrahan, S & Burgess-Limerick, R 2000. "Implementing Criterion-referenced Assessment within a Multi-disciplinary University Department." *Higher Education Research & Development*, vol. 19, no. 1, pp. 103-116.
3. Dorans, NJ 2002, *The Recentering of SAT® Scales and Its Effects on Score Distributions and Score Interpretations*, Research Report No. 2002-11, The College Board, viewed 2 May 2023, <https://files.eric.ed.gov/fulltext/ED563023.pdf>.
4. Jaynes, ET 2003, *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge, viewed 2 May 2023, <http://www.med.mcgill.ca/epidemiology/hanley/bios601/GaussianModel/JaynesProbabilityTheory.pdf>.
5. Lok, B, McNaught, C & Young, K 2016. "Criterion-referenced and Norm-referenced Assessments: Compatibility and Complementarity." *Assessment & Evaluation in Higher Education* vol. 41, no. 3, pp. 450-465, viewed 8 May 2023, <https://www.tandfonline.com/doi/full/10.1080/02602938.2015.1022136>.
6. Sadler, DR 2009. "Grade integrity and the representation of academic achievement." *Studies in Higher Education*, vol. 34, no. 7, pp. 807-826.
7. Baidoo-Anu, D & Owusu-Ansah, L 2023. "Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning." *SSRN*, 25 January 2023, viewed 8 May 2023, <https://ssrn.com/abstract=4337484 or http://dx.doi.org/10.2139/ssrn.4337484>.
8. Liao, L, Wong WH & Pang KM 2021. "An Action Research on Practicing Criterion-Referenced Assessment: The Case of General Education Foundation Programme of The Chinese University of Hong Kong." *Fudan Education Forum*, July 2021, vol. 19, no. 4. (This essay was written in Chinese.)
9. Cheung, HC, Chiu, CL, Gao, X, Liao, L, Pang, KM, Wong BW & Wong, WH 2018. "Criterion Referencing in GEFP: Practice, Review, and Vision." *EXPO 2018 Bulletin* (The Chinese University of Hong Kong), pp.151-166, viewed 8 May 2023, <http://www.cuhk.edu.hk/eLearning/expo/Expo2018-eBook.pdf>.

Appendix

Grade Boundary and Adherence to Weighting of Assessment Components

We consider a simple case where there are two assessment components. Below is a proof for mathematics amateurs.

Suppose that the score range for grade A and B are, respectively 80-100 and 61-79 and that there are two assessment components, Assignment 1 of weight x and Assignment 2 of weight $(1-x)$. (Hence the total weight is 1, that is, 100%. As a special case, if $x = 0.5 = 50\%$, the two components are of equal weight.) If a student gets mid-A and mid-B in Assignment 1 and Assignment 2 respectively, the final score (S) will be:

$$S = x\,(100 + 80)/2 + (1-x)\,(79+61)/2 = 70 + 20x \qquad (1)$$

If $x = 50\%$, the final score $S = 80$. To explain why a shift of a grade boundary would be equivalent to a change in the assessment scheme, we introduce a shift of the A to B boundary: A range is from $80+2p$ to 100 and then B range from 61 to $79+2p$, where p is a variable. For example, if A ranges from 85 to 100 and B from 61 to 84, p would be 2.5, the final score now becomes $S = x\,(100 + 80 + 2p)/2 + (1-x)\,(79+2p+61)/2 = 70 + 20x + p$. We rewrite it as

$$S = 70 + 20\,(x+p/20). \qquad (2)$$

By comparing Eq. (1) and Eq. (2), one can see that Eq. (2) is the total score the student gets if A ranges from 80 to 100 and B from 61 to 79, with the following assessment scheme: Assignment 1 weighs $x + p/20$ and Assignment 2 weighs $1 - (x + p/20)$. Hence, if two teachers use different score-grade conversion table in calculating students' final grades, a shift of a grade boundary will be equivalent to a change in the weights of each assessment components, or they are using different assessment schemes.