What are the Purpose and Properties of Normalization of Exam Marks for an Exam with Multiple Assessment Components (non-clinical and clinical)?

The transformation of exam scores through normalization of exam components may be understood through the following edited real question and answer.

Question:
Student A seems having a much higher score in SOE (Structured Oral Exam) which rates 30% while the total score is also higher. However, the transformed score is the reverse. Although quite marginal difference, it reverses their grade after transformation.

Initially, I thought the 30% rating in the SOE will help to pull up students who are good at the clinical part of the exam … It appears it is not enough …I wonder you would have any comment on this.

Student	Total	CA	SAQ	MCQ	SOE	Transf.Total	Grade
A	74.8	14.0	17.3	19.2	24.3	73.5	B+
B	74.0	14.7	19.0	21.1	19.3	74.1	A-

Class						
Max	100	20	25	25	30	100
Mean	70.2	14.8	16.6	19.0	19.9	70.2
SD	4.6	0.8	2.0	1.4	2.4	4.6

Answer:
To understand better the transformation of marks, please look at the scores in terms of percentile ranks:

Student	Total	CA	SAQ	MCQ	SOE	Transf.Total
A	88	19	50	53	100	63
B	75	53	97	97	44	78

Student A and B may represent the extremes of two type of students -- someone is good at Clinical and bad at book study vs. someone who is bad at Clinical while good at book study.

Student A has the highest ranking for SOE, being 100%, but all the book study works are mediocre to poor (about 50% or below). Student B has almost the highest ranking for SAQ and MCQ (97%) but being below average for SOE (44%).

What the transformation does is essentially, though not exactly (the ranks used are standardized marks), weighting the ranks instead of the raw marks of the students. In comparison, the student who has the higher weighted rank or weighted standardized scores (no. of SD from the mean) has the higher transformed scores. In terms of the non-clinical components, student B is much better than A in percentile ranking; while for clinical, student A is far better than B in rank. But since the non-clinical component is weighted 70% against the clinical of 30%, the students who are better in terms of rank in the non-clinical component will have the better transformed marks.

The transformation does not reverse the weighting preference of the exam components, but altering the final ranking a bit by considering the ranking instead of the raw marks of the students. This is done with the intention to pay justice to the different components that have different spreads, or SD (i.e. Components with a very narrow spread in marks are non-discriminative in raw marks for grading; that can be adjusted by using standardized marks or percentile ranks). To control for the differential significance of different exam components, the weightings of the component are the keys.

The transformation improves the discriminative power for all components with a relatively smaller SD, however, unfortunately, due to the majority weighting in the non-clinical components; the non-clinical

component exerts a more say in the discriminative ability. Hence the choice is to increase the weighting for the clinical components to the desired level.

Since the SAQ and MCQ combined have a weighting of 50%, the clinical SOE with 30% weighting cannot compete with the written exam in terms of influencing the final exam results. The only method that can be applied is to weigh the clinical component more than the non-clinical components. And even 40% in this case will not be enough. For example the clinical may be weighted as 50% (or more) and SAQ & MCQ combined as 40% and CA 10%. However, *in that case the reliability of the clinical assessment may become an important point to consider in mind as the clinical exam may tend to have fewer questions and be less reliable than the MCQs, or SAQ that usually have many more questions and by such having a higher reliability.*

Wing-Shing Chan, PhD
Office of Educational Services, CUHK (Tel: 2145-5234; Email: wschan@cuhk.edu.hk)