## WHY A PARADIGM SHIFT IN SCORING IS REQUIRED

With **scores** for testing and assessing, instructors usually mean numbers on some **bounded scale** between a given minimum and maximum score.

(1)

We will focus here on the **standard scale of percentages** from 0% to 100%, with *anchor points* 10%, 20%, ... , 90% in between.

Percentages are numbers between 0 and 1, e.g. 0.70, usually multiplied by 100 and then written as 70%. A percentage of 50% represents the **break**even score between passing and failing. Moreover, a score of 0% signifies the **highest degree** of failing and a score p = 100% signifies the **highest degree** of passing. So far, so good.

Common sense and fairness principles imply that the conventional rules of arithmetic cannot be applied to percentages. Below are some counterexamples dealing with inverses, sums and products of percentages, which clearly violate our intuition.

(3)

4	- 70% < 0%	70% + 20% = 90%	70% + 40% = 110%	2 * 70% = 140%	70% / 2 = 35%
	the inverse of 70% does not exist !	20% signifies failing, the sum should be < 70% !	the sum of percentages can't be > 100% !	the product of 70% with a number shall be ≤ 100% !	two failing scores of 35% lead to a passing score !?
5	e list of counter-intuitive	recults of When we ado	nt the rules of quasi-arithm		coring rule using groun-nee

applying ordinary arithmetic to percentages can be extended at will. Many **ad hoc solutions** have been proposed.

For example, the **capping operations min()** and **max()** are often applied, however, they are neither fair nor convenient to work with.

The weighted **arithmetic mean of scores** is often used, too, because it technically solves the problem of range violation. However, being based on ordinary arithmetic, it is just a *mean* trick, because it hides the other underlying **anomalies**.

Is the concept of percentage scores basically flawed and useless? Not at all. We need adequate operations! When we adopt the **rules of quasi-arithmetic**, developed over almost a century and applied in many areas of measuring, we can **avoid all problems** mentioned above, and work out a consistent and complete **scoring algebra**.

Actually, we only have to replace arithmetic addition by its **quasi-arithmetic** counterpart:

the **quasi-sum** of p and q is  $\left| \frac{p \cdot q}{p \cdot q + (1-p) \cdot (1-q)} \right|$ 

(2)

A quasi-sum lies between 0% and 100%. All

other **properties of addition** follow easily.

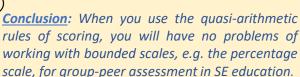
Moreover, **quasi-multiplication** of scores with any real number r can be easily constructed in terms of this new quasi-sum.

Finally, we get a well-behaved **inverse** for any score p, if we multiply it by -1. Then, we will have all we need for **quasi-arithmetic means** and **scoring rules**.

A quasi-arithmetic scoring rule using group-peer assessment may contain an impact parameter which moderates the effect which the peer ratings will have on the group score.

Moreover, scoring rules may also include the socalled **tolerance parameter** which restricts the final student scores to a specified subrange of the percentage scale around the group score.

Our **favorite quasi-arithmetic scoring rule**, with *impact* = 1 and *tolerance* = 2, is a simple function of student rating – with the group score and a calibration factor as parameters, such that all student scores will fall within a symmetric range around the group score.



(8)