Standard 500
From the “memoirs” of Cees Beuk

By Ed Zeelt, 2020

Beforehand

In 1983, as a newly appointed Superintendent of Vocational Education, I was asked what I thought of the International Professional Competitions. That was literally nothing, because I had never heard of it. But that year they were held in Linz (Austria), so I went there to look.

In Linz, I was introduced by René Gonthier (CH), chairman of the Technical Committee, by Arend Segaar, inspector of vocational education, in his position as Technical Delegate for the Netherlands. It turned out that Segaar, who knew me well from the Central LBO exams, had informed him in advance about my background: doctoral education and postdoctoral Methods / Techniques Social Science Education, research and in the leadership of the CITO with responsibility for test development and exams in vocational education and coordination at the Central Written Final Exams in Secondary Education.

Problem

This pre-information from Segaar had a clear intention. I had to be roped in for a job. Gonthier came directly to the point with the problem of experts who use only a small fraction of the 10-point grading scale and often almost only the numbers above five. High or low scores on the 100-point scale then say little about the level of the participants and, of course, it looks strange if the same score yields gold in one profession and you do not end up in the middle of the list in another profession. Therefore, the 100 point scale does not allow year-to-year comparisons to be made, or between countries, regions, etc.

Gonthier asked me for advice. I assured him that I had a relatively simple solution for making such comparisons. It would also have a positive effect on the comparability of the final figures, but that it might not be so easy to enter. “Count on the resistance of delegates and experts, who do not easily take away what they consider to be their competences. You have to invest in that. ”

Gonthier was not afraid of that. He was a school director. I could count on the support of a programmer from his school and the administrative help of his secretary.

Solution

My solution appealed to him. In short: all scores on the 100-point scale are linearly transformed into a scale with an equal (standard) value for all professions for an average score and one for all professions also with the same value for the distribution of the scores (standard deviation). A linear conversion does not change the ranking and distribution of scores. Nothing is added to or taken from the final assessment of the jury. That's exactly what they want but explain that.

Discussion paper

We agreed that I would put my solution on paper for consideration at a subsequent meeting of the Technical Committee. It included fine people but with little knowledge of statistics who should have little of it, so no formulas!

Definition of competition participants (competitors)

In my paper I defined the IVTO participants as a unique population of young champions in their field, who, after careful selection in previous rounds, have been selected at the top level for the International Professional Competitions (Skill Olympics) of the IVTO. It is therefore not a representative sample from an (imaginary) large collection of champions whose performances would show a normal distribution.

Standardize versus Normalize

Standardization is the linear conversion or transformation of scores. The ranking of the participants remains unchanged, as does the distribution of the scores. After standardization, the scores are still as far apart from each other as before conversion. The score distribution retains its shape during linear conversion (standardization); left or right skewed, flat, angular, etc. The (raw) scores are also not corrected for deviations from the normal distribution, because it is not available for the IVTO champions. Usually, standardization does not correct for klene groups.

This is often the case with normalization. Normalization refers to realized results from a group of test subjects representative of a particular population. If the distribution of the results of this research group deviates from the normal distribution, then this is due to this group and correction will take place.

Even with the normalization of the raw scores, the ranking of the participants on the new scale remains unchanged, but not the distribution of the scores. It will change. The raw scores are expressed somewhat plastically in the normal distribution model. One participant gets a little closer to its predecessor, while another takes a little longer. Expressively expressed "pieces of score scale" are exchanged between the contest participants. One hands in some of his score and another gets in. How much that is is difficult to say. This depends on the form of the score distribution. If it is close to a normal distribution, the pieces are small. With a strongly deviating shape, the pieces are large. But big or small, the ranking does not change, but for a participant it can make the difference between just sharing or not sharing a medal, or whether or not a diploma!

Standard scale

A standard score scale is a scale with no start (minimum score) or end (maximum score). There is only a mean (mean) for the average score of a group of participants, as well as a measure of the spread of the scores (standard deviation). Any value can be selected for both parameters.

My preference is for a standard scale with a mean of 500 points for the average score and 25 points for the spread (standard deviation).

A middle (mean) of 500 fits in well with the score and score scale used by the IVTO. An average score of 500 on the center of the standard scale logically follows a score of 50 points on the center of the score scale 10 -100 and the number 5 on the center of the grade scale 1 - 10. Although the meaning of a scale without beginning and end is completely different, they fit well together through their centers of 5, 50 and 500. Over the years, the 500/25 scale has proven to be useful in many (exam) situations and is generally accepted quickly.

Discussion

A mean of 400 or lower is of course also possible, but looks less attractive. It looks a bit like a scale for students with a fail. A center of 600 or higher resembles a scale for the top scorers.

I had no experience with those scales, but I wouldn't choose that. I found a 500-point scale to be a good fit with the IVTO's grading scale and raw scoring scale, partly because I didn't know what the effect would be on a jury.

Many juries are already with their scores and scores often above the number 5 on the score scale, and 50 points on the rough score scale, Renée Gonthier said. So they already use only half of the scale's range to differentiate. My fear was that at a higher midpoint of 600 or higher, they would further narrow the range of the grade and raw scoring scale in practice. There could then be a displacement of scores in the high part of the scale, resulting in unusually many medals to be shared.

Advice

My advice to the IVTO was to standardize the raw scores on a standard scale with a center (mean) of 500 points and a standard deviation of 25 or 30 points, without correcting anything and at least not normalizing the raw scores.

Trial

With the raw scores from IVTO's Youth Skill Olympics 1985 in Osaka, I tested the 500 scale with two standard deviations 25 and 30. I found no weird things, even with a few outliers (extremely high or low raw scores). The scale processed the scores without any problems. A standard deviation higher than 25 was not necessary to differentiate between the participants. That was enough at 25 points.

Cooperation

Gonthier, his programmer and I worked well together, and a nice computer program was created. But unfortunately, that led to a conflict between Daniel Sommer, the IVTO Secretary General and René Gonthier, which ran high and led to the departure of René Gonthier, his secretary and the program. It was a shame and a great pity because René was very competent, innovative and highly appreciated. I also discussed it with Francisco Albert, the President of the IVTO. He thought it was terrible but saw no solution. And so I was left empty-handed without a program and discovered that Gonthier had communicated absolutely nothing about it to Sommer. There were not even minutes of my discussions with the Technical Committee and other meetings. But the Technical Committee was fully aware of the plans, for which much support had since been gained. Daniel Sommer and his collaborator Erica Moser also pledged that support. And I knew another programmer who had written a program for me to study that we might use. The software had received a nice name from the programmer (BIC - Beech IVTO Calculations) and was certainly useful, but not really user-friendly. All data had to be entered manually. We did not have an optical reader available. In addition, in order to find out a little more about judging the individual jury member, the programmer had asked to standardize the sub-scores and to get the final scores from the average of the standardized sub-scores. I had not realized with this request that this would increase the manually typed data by a factor of 5 or 6! And in Sydney the program couldn't be changed for a while.

I am still deeply ashamed of this mistake, but when entering the data in Sydney, Erica Moser performed miracles together with Daniel Sommer. They worked much of the night to be able to present the results of the matches on the new 500 standard scale the same day. The introduction of the new scale 500 was a fact!

Erica Moser has completed the input of all data for the new 500 Scale - Sydney - 1988

The 29th International Skill Olympics took place at Darling Harbour Exhibition Centre in Sydney, Australia, from February 7th to 24th 1988.

FINAL RESULTS SYDNEY - 1988

	TRADES \ COUNTRIES	Austria	Australia	Bermudas	Brazil	Switzerland	Germany	Spain	Ireland	France	Liechtenstein	Gibraltar	Japan	Korea	Macao	Netherlands	New Zealand	Portugal	New Guinea	Taiwan	USA	TOTAL
1	Fitting	•	•		•	•	•					•	B	G	•	•	•	•	•	S		14
2	Press Tool Making	•	•			•	•		•			•	G	S		•	•			B	•	12
3	Instrument Making		•			B	•						•	G		•				S		7
5	Engineering Drawing		•		•	•	B		•		•		•	G	•		•		•	S	•	13
6	Turning	•	B		•	•	•		•			•	•	G		•	•	•	•	S	•	15
7	Milling		•		•	•	•		•			•	G	S		•		•	•	B	B	13
8	Constr. Steel Work		•			B	•		G	S			•	•		•	•		•	•	•	12
9	Gas Welding		•		•		•		G				•	G		•				B		8
10	Electric Welding		B		•		•		•				•	G		•	•	•	•	S	•	12
11	Wood Pattern Making	B	•			•	G					•	S	•		•	•			•		10
13	Panel Beating		•			•				•			B	G						S		6
14	Sheet Metal Work		•				•			•		•	B	G		•	•		•	S		10
15	Plumbing		S			•	•		•	•		•	•	G			•		•	B	•	12
16	Industrial Electronics	•	B	•	•	•	•		•			•	G	•	•					•	B	13
17	Consummer Electronics		•									•	•	G						S	B	6
18	House Wiring		•			•			•		•	•	G	S					•	B		9
19	Industrial Wiring	•	G			•	S		B		•	•	•	•			•	•		•	•	13
20	Bricklaying		G			•	B		•	S	•	•		•		•		•		•	•	12
21	Stonemasonry	•	S			•	•			•		G	B	•						•		9
22	Painting and Decorating	•	•				G		•			•	•	S		B						8
23	Plastering		•				•			G		S	•	B		•						7
24	Cabinet Making	•	•	•		S	•			B		•	•	•		•	•			G		12
25	Joinery	•	•			•	•		S	B	•	•	•	G		•	•	•		•		14
26	Carpentry		•			S	•		•	•		•	•	•		G	•		•	B		12
27	Jewellery	G	B									•	•	S		•	•					7
29	Ladies' Hairdressing	G	•				•					S		•		•	•	•		B	•	10
30	Mens' Hairdressing	G	•				•					S	•	•		•	B			•	•	10
31	Ladies' Dressmaking	B	•			•					•	B	•	G	•					S		9
32	CNC Machinery	S	•		•	•	B					•		G		•						8
33	Automobile Mechanics		S				B		•			G	•	•			•		•	•	•	10
34	Cookery	•	B	•		S	•		•		G	•	•	•	•	•	•				•	14
35	Waiting	G	•	•		•		B	•		S		•	•	•	•	•					12
36	Graphic Design		•	•									•	B						G	S	6
37	Agricultural Mechanics		G			S				B		•		•		•						6
	TOTAL	17	34	5	8	23	26	1	18	11	8	26	30	34	6	23	19	8	11	27	16	351

G - Gold S - Silver B - Bronze