Total Score of Athleticism: a strategy for assessing an athlete's athleticism

***Anthony Turner,** MSc, PGCE, CSCS*D, ASCC / London Sport Institute, Middlesex University*

OVERVIEW

The coaching staff that make up the multidisciplinary team of a sports club or governing body often require a single, holistic indication of an athlete´s athleticism. Currently there is no consensus on how this is best achieved and thus the Total Score of Athleticism (TSA) and its associated 'RAG' rating may provide one such strategy. The TSA is derived from the sum of z-scores from a fitness testing battery and is further divided into red, amber, green (RAG) zones based on the top (green) and bottom (red) third of all scores.

Introduction

As strength and conditioning coaches, we regularly put our athletes through a range of fitness tests to monitor progress and make adjustments to their programmes. Often, our results are also used to form a holistic opinion of an athlete by the sports club or governing body. In these instances, the psychologist, physiotherapist and the technical coaches also grade the athlete, with the results determining future interventions and team selection. Such situations lend themselves to providing a single score for the athlete's physical fitness, rather than, for example, separately discussing scores for jumps, speed or strength.

Also, these separate results can provide more information than those attending the meeting care need to hear. Currently, no method exists to provide a single score of holistic fitness, or rather athleticism, from a given fitness test and thus the aim of this paper is to suggest one such strategy.

The Total Score of Athleticism

The strategy in question is referred to as the Total Score of Athleticism (TSA) and is derived from the sum of z-scores. Outlined below is the calculation of the TSA using hypothetical data from a fitness testing battery; all calculations can be made using Excel and therefore this is available to most S&C coaches. Naturally, we need first to understand the z-scores before calculating and interpreting the TSA.

Understanding z-scores

It is prudent to start with an example of the z-scores' usefulness and although not entirely sports-related, Steinberg (2011) provides an excellent one. If an individual scores 72 on Test A and 84 on Test B, which did he do better on? The first stage of answering this is to establish what the possible points were. Test A may have been out of 95 points and Test B out of 105 points, leaving him with a performance of 76% and 80% respectively. So on a percentage basis, he now performed better on Test B. However, the 80% on Test B may have been one of the lowest scores amongst all test takers. Conversely the 76% may have been one of the highest, so arguably, he did best on Test A. Therefore, the final piece of information used is a measure of how well someone did relative to the others who took the test. A z-score gives us all this information.

Put in to a sports context, coaches and athletes are not as interested in the raw score as much as in where it ranked

AUTHOR'S BIO

ANTHONY TURNER, MSc, PGCE, ASCC, CSCS*D

Anthony Turner is the programme leader for the MSc in strength and conditioning at the London Sport Institute, Middlesex University. Anthony is the lead strength and conditioning coach for British Fencing, as well as a consultant to the Royal Army Physical Training Corps and also to various professional football clubs.

Table 1. Z-scores and the percentage of test scores they contain

Z-score	-3	-2.4	-2.1	-1.8	-1.5	-1.2	-0.9	-0.6	-0.3	0	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	3
%	0%	1%	2%	4%	7%	12%	18%	27%	38%	50%	62%	73%	82%	88%	93%	96%	98%	99%	100%

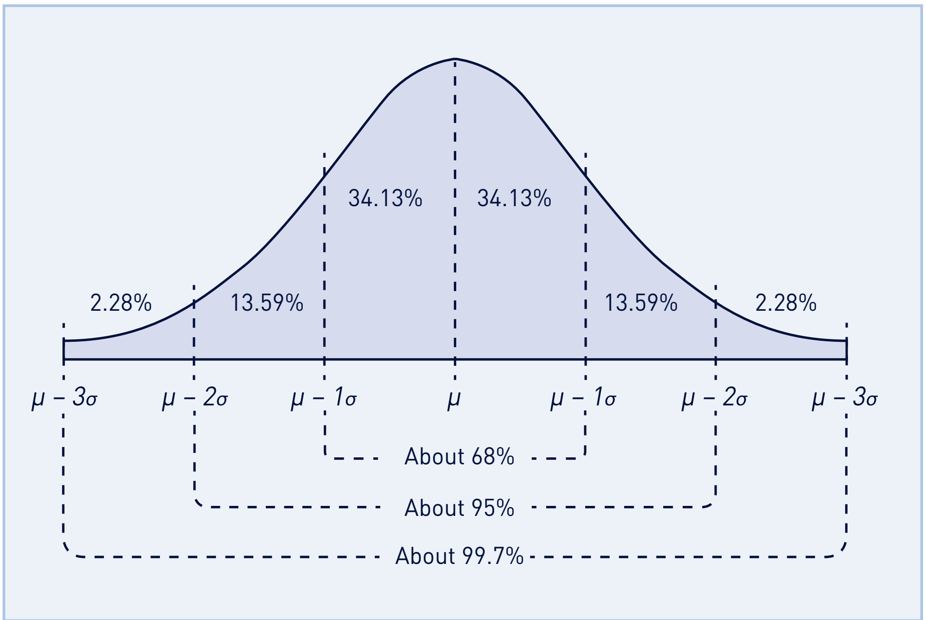


Figure 1. The mean ± the SD. The mean ± 1SD contains ~ 68% of all scores, ± 2SD ~95% and ± 3SD ~99%. μ = mean and σ = SD

amongst their team, especially when there is competition for places. For example, although norms derived from elite athletes are generally available for most fitness tests (and thus available for comparison), they may not be relevant for some of the following reasons:

1. **These elite athletes are not the ones competing for a place on this team!**
2. **They play the wrong sport – norms are not available for everything**
3. **This team is amateur or semi-professional, so using these benchmarks would establish unrealistic goals**
4. **The ultimate goal for most athletes is simply a place on the team, ie, they want to know who is the best amongst their peers.**

It should also be noted that it is unlikely that the person who achieved the highest bench press score will record, for example, the highest agility or Yo-Yo score, suggesting that there is some compromise among the different components of athleticism that make up a good sports performer. So, although the z-score provides 'teammate specific' scores and benchmarks, as well as highlighting weaknesses amongst players relative to their team mates, the TSA (ie, the sum of z-scores) also uses these scores in a way that recognizes their independence from each other, but also recognizes the significance of having high scores for each.

Using the mean and standard deviation

To understand z-scores further and interpret the data, it is important to now look at the mean and standard deviation (SD). The former provides the average score and the latter the dispersion of data from the mean; both are used to rank the data amongst the given population (as described above). Together:

- **The mean ± 1SD will contain ~ 68% of all test scores,**
- **The mean ± 2SD ~ 95% and**
- **The mean ± 3SD ~ 99% (see Figure 1 and Table 1).**

For a z-score, all data is converted to have a mean of zero and a SD of one. So, if an athlete scores +2, it indicates that the athlete scored above the mean by 2SD. This corresponds to the athlete performing better than 97% of all scores (50% up to the mean plus 34% up to +1SD and another 13% up to +2SD; see Figure 1 and Table 1). A score of +1 informs us that they scored better than 84% of others who were tested; conversely -1 suggests 84% did better than them. Of note, in sport, smaller values can be a sign of better performances; for example, 30m-sprint time. Here, negative values for z-scores would be produced for athletes who were better than average.

Where this occurs, and especially importantly when histograms or radar plots are produced to highlight athlete strengths and weaknesses (see Figure 2), the final value can simply be multiplied by -1 (discussed later). The equation for calculating the z-score is outlined below (Equation 1). Because it uses the SD, the scores are unitless, thus enabling the summing of results across all tests. An additional benefit of using z-scores is that graphing the data (as in Figure 2) clearly highlights an athlete's strengths and weaknesses across fitness tests relative to their teammates. These can be useful in guiding training programme design.

Using excel to calculate z-scores

The data is first entered with one athlete per row, with tests in columns across the top (see Figure 3). The mean for each test should be calculated and in Excel is done using the formula **'=average(cell range)'**. The SD should then be calculated using **'=stdev(cell range)'**; Figures 3 and 4 illustrate this process, with the formula bar containing the calculation, included at the top of each figure. Once this has been done for all tests, z-scores can be calculated: this is illustrated in Figure 5. The z-score formula in excel is **'=(raw score cell – X cell)/SD cell'**. Notice that in Figure 5 (see formula bar at top), **'$'** symbols are used between the column letter and cell number for both the mean and the

SD, i.e., ‘**B$15**’ and ‘**B$16**’ respectively. The ‘**$**’ symbol stops the row moving as the formula is dragged down to cover all athlete scores. Also, in this example, the z-score equation is bracketed and then multiplied by -1. As discussed previously, this is because we actually want athletes to achieve lower scores for timed tests of this kind, and thus it helps with graph interpretation. Of course, this should not be done for jump test scores.

Calculating the Total Score of Athleticism and ‘RAG’ rating them

Once the z-scores have been calculated, they can simply be summed to calculate the TSA (see Figure 6a) and again graphed to further highlight athlete athleticism relative to their teammates (see Figure 6b). TSA scores can then be ranked and colour coded to represent a ‘RAG’ rating. ‘RAG’ refers to red, amber, green and is a way of colour coding the athletes in to a ‘safe’ (green), ‘caution’ (amber) and ‘danger’ zone (red); of course, athletes in the ‘red’ need to improve test performance as their team selection is negatively affected by this. Such coding again helps in providing each value with a quick reference indication of athletic performance.

To rank the athletes use the function: **‘=rank(the number you are trying to rank, the cells from which the rank is derived)’**. Figure 7 illustrates this in the formula bar and again note the use of ‘**$**’, so that the formula can be dragged down. Finally, the ranks can be RAG rated using the ‘**conditional formatting**’ option as illustrated in Figure 8. Here, the ‘**traffic light**’ option is selected as scores are automatically split into thirds. However, it is important to go back into the conditional formatting, click on the ‘**manage rules**’ tab, then the ‘**edit rule**’ option and finally check the ‘**reverse icon order**’ box, so as the order of traffic lights can be reversed. That is, the score ‘1’ is ranked the best and indicated by a green traffic light.

Which tests should be used?

It is beyond the scope of this paper to fully discuss how to construct a fitness testing battery; this has been discussed previously in tennis (Roberts et al, 2011), rugby league (Coneyworth et al, 2012), soccer (Turner et al, 2011), fencing (Turner et al, 2013), taekwondo (Turner et al, 2009) and ice hockey (Nightingale et al, 2013) for example. Readers are recommended to read these papers for further information, noting that the principles that underpin the construction of fitness tests within these sports can be applied across others.

It is important to choose the right tests as the TSA is a summation of these scores; therefore the battery, or selected tests to be used within the TSA, largely governs the validity of this method. The general consensus is to test strength, power, speed, agility, speed-endurance and aerobic capacity within a test, although not all of these will be required and the way in which these are tested will differ depending on the sport and the equipment available. For example, for some sports such as fencing, aerobic capacity may not be relevant. Speed, via traditional linear sprinting, is not necessary in taekwondo and although it is tested over 10, 20 and 30m in soccer, it would be invalid to test it over 10m in tennis. Agility in soccer may be best tested using the pro-agility test, and in rugby league via the L-run.

The one thing all these tests have in common is the need to test strength and power. These can be done in the gym (eg, 1RM back squat, 1RM power clean) but requires specialist equipment and for athletes to be technically sound at performing them. Where both these criteria are not met, a squat jump and countermovement jump (respectively) can be used as a ‘field-based’ alternative or even additions to gym-based tests.

In summary, the tests need to contribute to performance: if there is no obvious relevance in terms of movement patterns or force generation characteristics, the test may be invalid. The importance of using the correct tests cannot be understated.

The TSA and its comparison to the ‘EPIC rating’

The TSA draws parallels with a testing battery and subsequent ‘EPIC rating’ devised by Boyd Epley at the University of Nebraska. Here, performance over a 10-yard dash, vertical jump, pro-agility and the hang clean are used via an algorithm scaled for body mass, to produce a single score of athletic talent; scores range from 0-1000. The tests and scoring system used appear to be born from an analysis of over 20,000 Division One American college athletes.

Regretfully, there is not much more information available (to the author´s knowledge) on this protocol, despite it being endorsed by the National Strength and Conditioning Association back in 2010. Although it is a promising tool, the tests are predetermined and in many instances, absolute rather than relative scores may be preferred. One of the benefits of the EPIC rating, however,

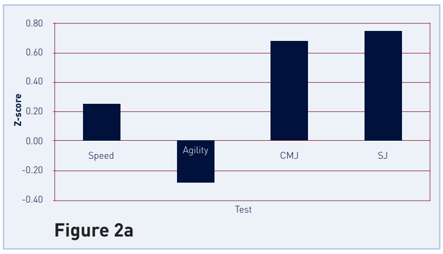


Figure 2a

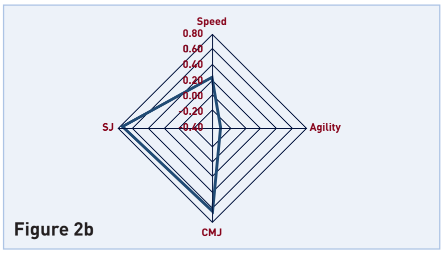


Figure 2b

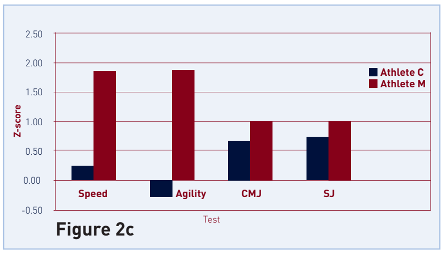


Figure 2c

Figure 2a-c. (left to right). Example results, illustrated as z-scores, produced by an athlete as part of squad fitness testing (CMJ = countermovement jump; SJ = squat jump). Here (Figure 2a) the athlete is above the squad average (represented by the zero line) for speed and jump tests but not agility (which should therefore be targeted in subsequent programming). This graph can also be presented as a radar plot if preferred (Figure 2b) or used to compare athletes (Figure 2c). Equation 1 z-score = (X-M)/SD
Where X = the raw score, M = the mean and SD = standard deviation

B15 | =AVERAGE(B2:B14)

	A	B	C	D	E
1	Athlete	Speed	Agility	CMJ	SJ
2	A	4.04	5.09	49.90	41.80
3	B	4.20	4.89	41.30	36.00
4	C	4.07	5.05	56.50	50.10
5	D	4.11	5.43	53.70	45.20
6	E	4.04	4.92	56.60	53.30
7	F	4.26	4.97	47.10	43.30
8	G	4.05	4.74	48.50	44.20
9	H	3.92	4.55	66.00	56.80
10	I	4.15	4.93	52.80	44.40
11	J	4.28	4.84	47.70	41.10
12	K	4.09	4.94	47.60	39.60
13	L	4.23	5.70	52.80	46.90
14	M	3.87	4.28	58.70	51.60
15	Mean	4.10	4.95	52.25	45.72
16	SD	0.12	0.35	6.32	5.88

Figure 3. Data input for a fitness testing battery, with the calculation of test score means illustrated in the formula bar

B16 | =STDEV(B2:B14)

	A	B	C	D	E
1	Athlete	Speed	Agility	CMJ	SJ
2	A	4.04	5.09	49.90	41.80
3	B	4.20	4.89	41.30	36.00
4	C	4.07	5.05	56.50	50.10
5	D	4.11	5.43	53.70	45.20
6	E	4.04	4.92	56.60	53.30
7	F	4.26	4.97	47.10	43.30
8	G	4.05	4.74	48.50	44.20
9	H	3.92	4.55	66.00	56.80
10	I	4.15	4.93	52.80	44.40
11	J	4.28	4.84	47.70	41.10
12	K	4.09	4.94	47.60	39.60
13	L	4.23	5.70	52.80	46.90
14	M	3.87	4.28	58.70	51.60
15	Mean	4.10	4.95	52.25	45.72
16	SD	0.12	0.35	6.32	5.88

Figure 4. How to calculate test score standard deviation (SD), with the formula illustrated in the formula bar

is that you can track the athlete's score over time, with higher scores reflecting better athletic potential.

In contrast, because the TSA is derived from z-scores, an athlete's athletic ability is always judged relative to his teammates and so fluctuations can represent changes in the athlete, the teammates, or both. Therefore, when tracking changes, the raw score should be used but of course, the athlete should always aim to be ranked number one for athleticism, and be in the green.

Conclusion

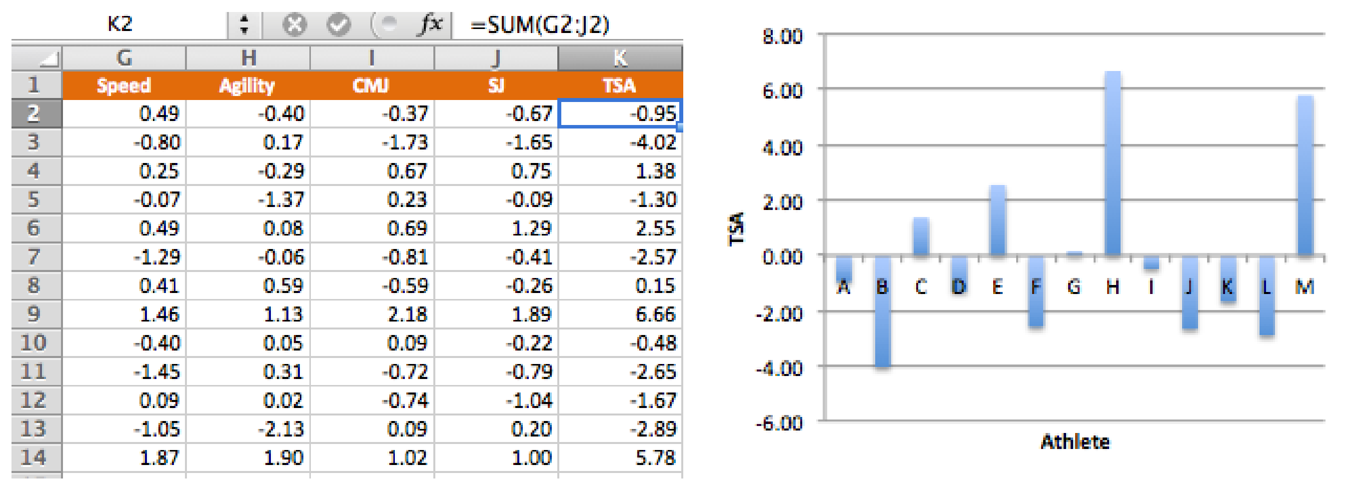
Often the coaching staff that make up the multidisciplinary team of a sports club or governing body require a single, holistic indication of an athlete's athleticism. Currently there is no consensus on how this is best achieved and thus the TSA and its associated 'RAG' rating may provide one such strategy. The validity of the TSA score is largely determined by the relevance of the fitness tests used, so coaches must be able to rationalise their choices.

Finally, the z-score graphs can also be used to guide the athletes' training programme. The figure schematic is such that it is relatively simple to interpret for both coaches and athletes. For histograms, for example, above the line means an athlete is better than average, while below the line suggests they are worse. A score of 1 indicates an athlete scored better than 84% of his teammates, while a score of 2 indicates he scored better than 97%.

G2 | =((B2-B$15)/B$16)*-1

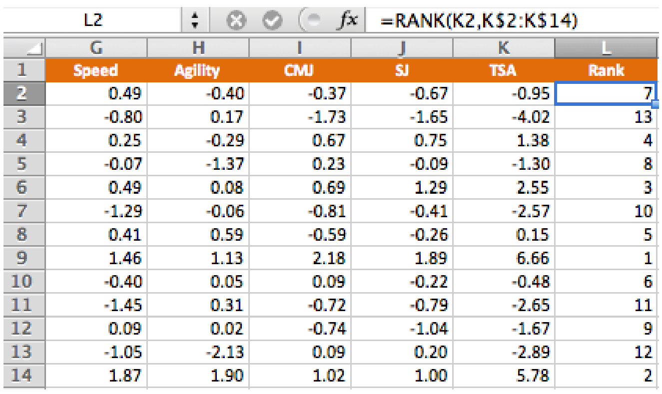
	A	B	C	D	E	F	G	H	I	J
1	Athlete	Speed	Agility	CMJ	SJ		Speed	Agility	CMJ	SJ
2	A	4.04	5.09	49.90	41.80		0.49	-0.40	-0.37	-0.67
3	B	4.20	4.89	41.30	36.00		-0.80	0.17	-1.73	-1.65
4	C	4.07	5.05	56.50	50.10		0.25	-0.29	0.67	0.75
5	D	4.11	5.43	53.70	45.20		-0.07	-1.37	0.23	-0.09
6	E	4.04	4.92	56.60	53.30		0.49	0.08	0.69	1.29
7	F	4.26	4.97	47.10	43.30		-1.29	-0.06	-0.81	-0.41
8	G	4.05	4.74	48.50	44.20		0.41	0.59	-0.59	-0.26
9	H	3.92	4.55	66.00	56.80		1.46	1.13	2.18	1.89
10	I	4.15	4.93	52.80	44.40		-0.40	0.05	0.09	-0.22
11	J	4.28	4.84	47.70	41.10		-1.45	0.31	-0.72	-0.79
12	K	4.09	4.94	47.60	39.60		0.09	0.02	-0.74	-1.04
13	L	4.23	5.70	52.80	46.90		-1.05	-2.13	0.09	0.20
14	M	3.87	4.28	58.70	51.60		1.87	1.90	1.02	1.00
15	Mean	4.10	4.95	52.25	45.72					
16	SD	0.12	0.35	6.32	5.88					

Figure 5. How to calculate test score z-scores (see highlighted cell G2), with the formula illustrated in the formula bar



	G	H	I	J	K
1	Speed	Agility	CMJ	SJ	TSA
2	0.49	-0.40	-0.37	-0.67	-0.95
3	-0.80	0.17	-1.73	-1.65	-4.02
4	0.25	-0.29	0.67	0.75	1.38
5	-0.07	-1.37	0.23	-0.09	-1.30
6	0.49	0.08	0.69	1.29	2.55
7	-1.29	-0.06	-0.81	-0.41	-2.57
8	0.41	0.59	-0.59	-0.26	0.15
9	1.46	1.13	2.18	1.89	6.66
10	-0.40	0.05	0.09	-0.22	-0.48
11	-1.45	0.31	-0.72	-0.79	-2.65
12	0.09	0.02	-0.74	-1.04	-1.67
13	-1.05	-2.13	0.09	0.20	-2.89
14	1.87	1.90	1.02	1.00	5.78

Figure 6a and b. (left to right). The Total score of Athleticism (TSA) is calculated by adding up the z-scores of each athlete; the formula for this is illustrated in the formula bar (Figure 6a). The athletes can then be ranked according to this score (see last column). The TSA can then be compared across athletes as in Figure 6b.



	G	H	I	J	K	L
1	Speed	Agility	CMJ	SJ	TSA	Rank
2	0.49	-0.40	-0.37	-0.67	-0.95	7
3	-0.80	0.17	-1.73	-1.65	-4.02	13
4	0.25	-0.29	0.67	0.75	1.38	4
5	-0.07	-1.37	0.23	-0.09	-1.30	8
6	0.49	0.08	0.69	1.29	2.55	3
7	-1.29	-0.06	-0.81	-0.41	-2.57	10
8	0.41	0.59	-0.59	-0.26	0.15	5
9	1.46	1.13	2.18	1.89	6.66	1
10	-0.40	0.05	0.09	-0.22	-0.48	6
11	-1.45	0.31	-0.72	-0.79	-2.65	11
12	0.09	0.02	-0.74	-1.04	-1.67	9
13	-1.05	-2.13	0.09	0.20	-2.89	12
14	1.87	1.90	1.02	1.00	5.78	2

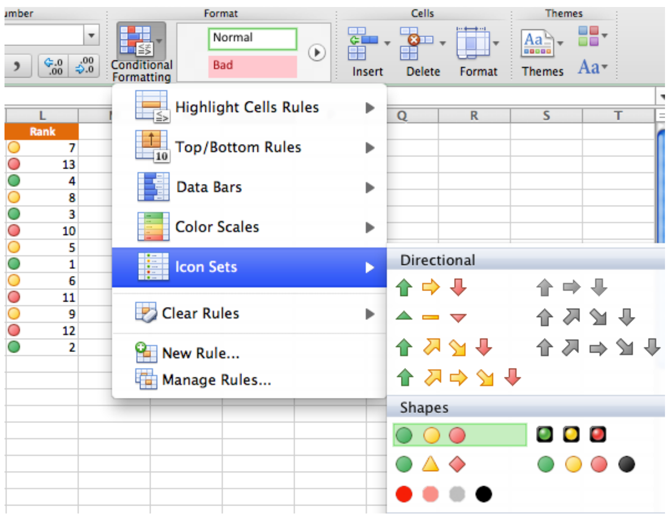


Figure 7. To rank the scores for TSA, use the function highlighted in the formula box and use the '$' so it can be dragged down for all scores

Figure 8. Use the 'conditional formatting' option to traffic light the data into thirds. However, you will need to edit this rule so a score of '1' is the highest and gets a green light

References

1. Coneyworth, P., Turner, A., & Ward, N. A field based test battery for rugby league. *Professional Strength and Conditioning* , 22, 6-12. 2012.
2. Nightingale, P, Miller, S., and Turner, A. The usefulness and reliability of fitness testing protocols for ice hockey players: a literature review. *Journal of Strength and Conditioning Research*, 27(6): 1742-1748. 2013
3. Roberts, T., Turner, A., & Mayers, R. A field based test battery for tennis. *Professional Strength and Conditioning*, 20, 13-22. 2011.
4. Steinberg, WJ. *Statistics alive!* 2nd edition. Thusand Oaks CA: Sage publications. 2011
5. Turner, A., Miller, S., Stewart P., Cree, J., Ingram, I., Dimitriou, L., Moody, J., and Kilduff, L. Strength and Conditioning for Fencing. *Strength and Conditioning Journal*. 35 (1), 1-9. 2013.
6. Turner, A. Strength and conditioning for taekwondo athletes. *Professional Strength and Conditioning* , 15, 15-27. 2009.
7. Turner, A., Walker, S., Coneyworth, P., Reed, G., Barter, P., Hembridge, M., et al. A field-based testing battery for the assessment of fitness in soccer players. *Strength and Conditioning Journal*, 33 (5), 29-39. 2011.