**The role of analytics in assessing playing talent**

**References:**

<https://eprints.whiterose.ac.uk/98808/3/Book%2520Chapter%2520Role%2520of%2520Analytics%2520in%2520Assessing%2520Playing%2520Talent%2520%28Final%2520Feb%25202016%29%5B1%5D.pdf>

**Abstract:**

In many invasion-territorial sports, player performance is multi-dimensional with individual players required to undertake a variety of different actions in offense and defence. The degree of specialisation by individual players varies across sports with association football (i.e., football) towards the “generalist” end of the spectrum while American (gridiron) football is more towards the “specialist” end of the spectrum. But, irrespective of the degree of individual specialisation, it follows that the invasion-territorial sports require an array of performance metrics to capture the multidimensional skill sets that individual players and teams require. Any summary player performance metric must be of necessity a composite player rating that is based on combining a set of skill-specific performance metrics. Unlike baseball there is no single skill-specific performance metric that can effectively capture an individual player’s contribution to team performance and game outcome. So, player rating systems in the invasion-territorial sports necessarily comprise two distinct problems, an identification problem of determining the most appropriate set of skill-specific performance metrics, and a composition (or weighting) problem of how to best combine the set of skill-specific performance metrics to construct a summary player performance rating. This goes way beyond Moneyball which focuses mainly on the identification problem in the context of finding the single best metrics for the two core 11 skills in baseball, pitching and batting. (Statistically, fielding is of minor importance as a systematic determinant of game outcomes.)

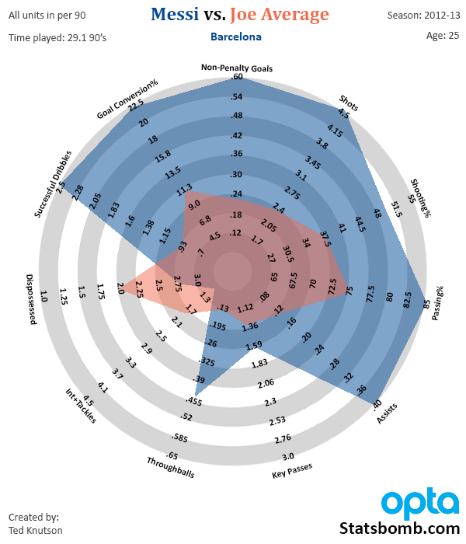
**Expert Judgment versus Statistical Analysis:**

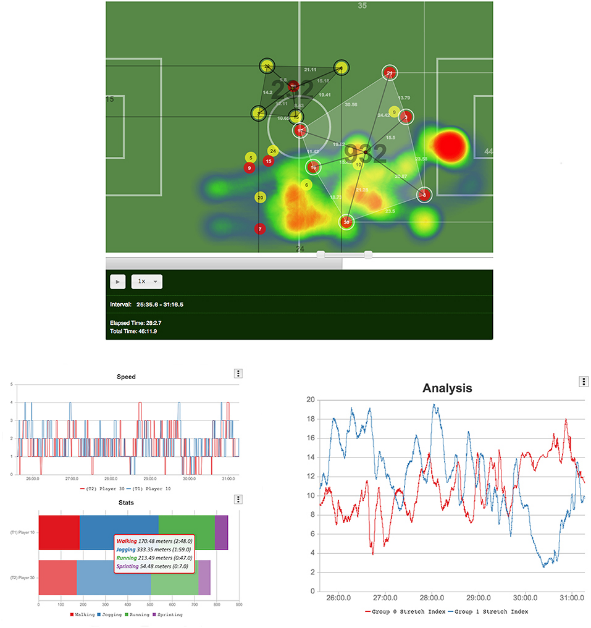
The issues facing the development of a more analytical approach to talent ID and player recruitment decisions in the invasion-territorial team sports are not unique. Indeed, there is a large body of decision research on the relative merits of expert judgment and statistical analysis as the basis for decisions on the best course of action in multivariate contexts. Dawes (1988) actually tracks the antecedents of this research right back to Benjamin Franklin in the 18th Century and Franklin’s proposal of the method of “prudential algebra” in which the reasons for and against each alternative course of action are identified and assigned a score of +1 or -1, respectively, with the recommended course of action having the highest net score. The findings of this research particularly over the last 60 years are remarkably consistent and very instructive for the specifics of how to most effectively utilise the contributions of coaches, scouts and data analysts. The starting point for the modern research on expert judgment and statistical analysis as the basis for decision making is Paul Meehl’s book, Clinical versus Statistical Predictions: A Theoretical Analysis and Revision of the Literature published in 1954. Meehl compared the findings of 20 different studies in a wide range of areas and discovered that statistical analysis always provided at least as good predictions of future outcomes, and in most cases significantly more accurate predictions, than the predictions of experts using their intuition and experience. Meehl’s book, which he himself described as “my disturbing little book”, provoked considerable controversy at the time and led to further studies comparing the effectiveness of experts and algorithms. This research continues. But the overwhelming body of evidence points in one direction, namely, the superiority of algorithms over experts. When it comes to man versus machines in predicting the outcomes of different courses of action, it is as close to a unanimous verdict as could be expected in the real world.

**Player Rating Systems with Multiple Performance Metrics:**

As previously discussed, as soon as we move into the invasion-territorial team sports, there is a need to develop player rating systems that combine multiple performance metrics. Applying the findings from decision research, most prominently Meehl, Dawes and Kahneman, the key argument for using player rating systems is that they ensure consistency in the comparison of all players. The actual weightings used to combine the multiple performance metrics into an overall player rating is a secondary concern.

Improper linear models especially equal weights models are often as accurate in their predictions as proper linear (i.e., regression) models. Using equal weights ensures consistency but avoids using weights that have been derived from one specific sample. regression coefficients play another role apart from optimising the relative importance of individual predictors. Regression coefficients also control for differences in the units of measurement across the predictors. This is an important consideration in player rating systems particularly when often the individual performance metrics are of two broad types with very different units of measurement. Performance metrics can often be categorised as either activity levels or success ratios. Activity levels are tally counts of the frequency with which a player has performed a specific action such as the number of attempted passes, the number of attempted tackles and the number of shots at goal. By contrast, success rates show the proportion of successful outcomes relative to the total number of attempts and are often reported as percentages. Pass completion, tackle success and shot accuracy are all examples of success ratios. Given the very different units of measurement involved in activity levels and success ratios, it follows that any useful player rating system must allow for these measurement differences as well as taking account of relative importance considerations. So, even if the equal-weights approach is adopted to develop a player ratings system that combines a number of skill-specific performance metrics, this is insufficient on its own. The skill-specific performance metrics need to be standardised before being combined into a composite player rating. The most frequently used method of standardisation is Z-scores in which each performance metric is expressed as a deviation from its mean value divided by its standard deviation. The implications for talent ID and assessing the development of youth players are very clear. An effective player rating system must be comprehensive which requires that it includes metrics that capture all of the relevant factors and these metrics must be combined in a consistent manner. The expertise of coaches and scouts is the crucial starting point for determining the relevant factors that are able to identify young talented players with the highest success probabilities of a career at the elite level of their sport. In this discovery phase, the analysts play a secondary role in formulating the precise metrics to measure the relevant factors as well as validating the degree to which these metrics are predictive of future career success. Having agreed the set of metrics to be used to identify young talent players and track their development progression, the analyst then has the task of developing a composite rating that brings together all of the metrics. The findings of decision research suggest that the rating system does not necessitate the use of sophisticated statistical multivariate techniques such as multiple regression.

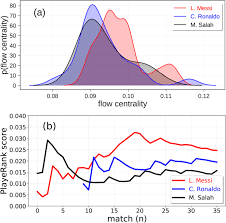




**Overcoming the Clash of Cultures in Elite Sports:**

The clash of cultures, art versus science, man versus machine, creates the dramatic tension that runs throughout the movie with analytics ultimately winning the day, captured in the moment when the home run that seals the A’s record-breaking 20-game winning streak is scored by a player recruited on the basis of his statistics in the face of opposition from the scouts. The contrast can also be seen, for example in the anti-analytics baseball movie, Trouble with the Curve, starring Clint Eastwood in which traditional scouting triumphs by detecting a fundamental flaw in the batting technique of a first-round draft pick with great metrics as well as discovering a great pitching prospect for whom there is no data and so would never register on the analyst’s radar.

Coach-led analytics utilises the coach (or scout) as the expert best able to identify a comprehensive set of player characteristics that predict future performance. But what is evident is that expertise in the identification problem does not translate into expertise in the composition problem. Indeed, expertise in identifying performance predictors often gets in the way of applying consistency in combining these performance predictors. Coaches and scouts have a tendency towards inconsistency by rating individual players individually, giving more weight to certain predictors for some players but not all. The evidence clearly shows that this subjective and selective application of algorithms is likely to diminish rather than enhance the effectiveness of the decision-making process. Hence the key lesson from decision research that algorithms whether statistically-based or applying equal weights to standardised metrics are the best way to support decision makers. Effective decision making is an art and a science. Analytics has a key role to play in supporting coaching decisions but so too has the experience and expert judgment of the coaches and scouts. The most successful teams are likely to be those that can combine effectively both sources of input into the decision calculus.



**My critiques and improvements:**

In my opinion, as has been prefaced before analytics in football is heavily dependent on who is performing the identification and analytics. It is also dependent on the quality and the availability of data. My improvement would be improving the collection of data around the world but also using intangible metrics to perform analytics in order to identify future prospects and recruitment.

**Conclusion:**

The increased usage of data and player statistics will be a part of the recruitment space for years to come. Extremely talented scouts will be challenging to find hence are of great value to any club.Both statistics and traditional scouting will have to go hand in hand to identify the best player. Scouting Biases are common but must be eradicated so that talented players are given a chance to prove themselves. Using Machine Learning and [Artificial Intelligence](https://www.latentview.com/glossary/#A) for predictive analysis and creating new applications has increased recently. This can be done by analysing data and gaining accurate insights. **– Siddharth Nair (PES2UG20CS339)**