- ○ Longform
- ○ Cycling
- ○ Boxing
- ○ Horse Racing
- ○ Podcasts
- ○ Local
- ○ Running
- ○ Water Polo
- ○ Wrestling
- ○ Outsports
- ○ Video
- ○ Odds

- All 308 blogs
-

-
  Follow
-
  +1 1.9M
- Search

  Search

- NHL
- Outnumbered

The tricky thing about finding the 'be all, end all' stat

By Eric T. @BSH_EricT on Jan 13 2014, 12:27p 7

Justin K. Aller

Measuring a player's value with a single number is a worthy pursuit, but it's a really tough thing to do in hockey. Here's one of the reasons why.

⇅ Tweet Share on Twitter (19) ⬆Share Share on Facebook (5) ⬆ Share Share on Google Plus 7 Comments ⋆Rec Recommend this Post

One question I often get asked is why there's no hockey equivalent of WAR, a single metric that we can use to encapsulate a player's value.

There actually are a bunch of value metrics out there -- GVT, DeltaSOT, HART, THoR, and several others. But instead of getting into the strengths and weaknesses of each, I usually answer the question with a more general comment:

More *Outnumbered*

I think we have reasonable measures of many aspects of play, but that we don't know enough yet about how the pieces fit together to rely on a simple formula to integrate them. So I prefer to force people to think about the pieces by leaving them un-integrated.

It's not as user-friendly, but I think it leads to more trustworthy analysis. It's too easy to blindly use a value metric without thinking about its weaknesses -- I almost never hear someone say "GVT says this player is worth 6 goals per year, but we know its formula will overrate his defense because..."

The whole reason people want a comprehensive value metric is so they don't have to make those caveats, so they can just use a simple number. Until we can make something reliable enough to be worthy of that trust, I'd rather not lead people down that path.

That's not as easy as you might think.

A tricky problem

Obviously, I wouldn't study analytics if I didn't think there were things they can do better than people's observations. But analytics do have their weaknesses, of course. Let's talk about one of the issues that I think is particularly difficult for a computer and fairly easy for a human.

Suppose there was a team with a roster that looked something like this (I use the word "skill" here to include all abilities and intangibles that might affect the outcome of a game):

Players	Skill level compared to average F/D
First line	20 goals better than average
Second line	10 goals better than average
Third line	5 goals better than average
Fourth line	5 goals worse than average
First pair	5 goals better than average
Second pair	15 goals worse than average
Third pair	20 goals worse than average

Add it all up and you'll find that it adds up to 0 -- the team that we're talking about is average.

Remember that nobody is omniscient, so we don't know these underlying skill levels. All we know is the team's actual results. So what would we observe for the year?

The forwards are collectively worth 30 goals more than average. So we can roughly say they boost each defensive pairing's results by 10 goals. (This won't be exactly right if the pairs get different ice time -- accounting for that will make the math more complicated but won't change the result).

So with that 10-goal boost from the strength of the forwards, the first pair would be +15, the second pair would be -5, and the third pair would be -10. Remember that, we'll come back to it.

Let's imagine a second team. This team has a roster that looks like this:

Players	Skill level compared to average F/D
First line	10 goals better than average
Second line	5 goals worse than average
Third line	15 goals worse than average
Fourth line	20 goals worse than average
First pair	25 goals better than average
Second pair	5 goals better than average
Third pair	Average

This team is average too -- their total is also 0. And what do we observe for their results?

Now our forward group collectively has a -30 goal talent, so they bring each defensive pairing down by 10 goals. So our first pair finishes the year at +15, our second pair is -5, and our third pair is -10. Sound familiar?

Analytically, differentiating between these two teams is pretty close to intractable. We can tell that the first pair is much better than the rest of the defensemen, but we can't tell whether they're a premium pair being compared to an average group or whether they're an average pair and the guys behind them are losers.

That's because the computer can only tell that overall the team is average; it can't tell whether the forwards are collectively better or worse than the defensemen. And that's something that's really not hard at all for a human to do.

Getting closer

A good algorithm could start to address the problem by paying attention to what happens when a player changes teams. If a guy is the best defenseman on one team and just looks like one of the guys on a new team, the program might infer that the second team has a stronger, deeper group of defensemen.

But occasionally, the issue will be that the player was playing through an injury in the second year, or just ran cold for a while for whatever reason. And since there isn't a ton of player movement, we'll be relying pretty heavily on those few players to help us sort this out.

As the analysis gets more and more sophisticated, we'll chip away at issues like this, and a value metric will become more reliable.

But given where we are right now, I don't trust our algorithms to evaluate players without human supervision; I'm more comfortable giving people the tools and letting them piece together a picture using both the numbers and what they know about the players.

So while I think teams are making a mistake if they try to run their organization without help from the computers, I don't think the computers would do a very good job of running the organization without people either.

More from SB Nation NHL: