

The Base Advancement Percentage

by Richard T. Newell Jr.

The theory behind the base advancement percentage (BAP) is that the more bases a batter can advance himself and any runners on base when he is at the plate, the more he is helping his team. The BAP attempts to measure how successful a batter is in doing that by determining what percentage of potential base advancement the batter accomplishes. In its simplest form, it would be calculated as follows:

$$\text{BAP} = \frac{[\text{Number of bases advanced by the batter and runners during the batter's at bat} - \text{Number of bases lost during the batter's at bat due to the out of a runner}]}{\text{Possible number of bases that could be advanced by the batter and runners when the batter comes to the plate}}$$

The following examples show the calculation of the BAP:

Assume the batter comes to the plate with a runner on first base. The possible number of bases that could be advanced during the at bat is 7 (3 by the runner on first and 4 by the batter).

Example Result 1: The batter hits a home run. The result is an advancement of 7 bases (i.e., the runner advances 3 bases and the batter 4). $\text{BAP} = 7/7 = 1.000$

Example Result 2: The batter singles and the runner advances to third. The result is an advancement of 3 bases (i.e., the runner advances 2 bases and the batter 1). $\text{BAP} = 3/7 = .429$

Example Result 3: The batter pops out, and the runner holds. The result is an advancement of 0 bases. $BAP = 0/7 = .000$

Example Result 4: The batter grounds into a double play. The result is an advancement of -1 bases (the runner loses the 1 base he had previously advanced). $BAP = -1/7 = -.143$

As can be seen by the examples above, the maximum attainable BAP is 1.000, while a negative BAP is possible. (The worst possible BAP as defined above would be $-.714$, achieved when a batter comes to the plate with runners on second and third and then lines into a triple play. However, using a modified version suggested below where only runners' outs due to force outs create negative base advancement figures, the worst possible BAP would be limited to $-.500$.)

A BAP can be calculated for each at bat. The results of each at bat (i.e., the actual base advancements and the possible base advancements) can be cumulated to obtain a cumulative BAP.

An advantage of the BAP relative to more often used hitting statistics like batting average and slugging percentage is that it gives greater weight to at bats where there are runners on base. As such, one might argue that it incorporates to some extent how a batter performs in "clutch" situations. It also gives some reward for doing some of the little things (e.g., hitting behind a runner so the runner can advance an extra base) that are valued in baseball but not picked up in the usual statistics. And, like slugging percentage, the BAP gives a higher reward for producing extra-base hits than for singles.

Questions and Possible Variations

What happens when any base runner's situation changes in the middle of an at bat (e.g., a runner steals a base or is thrown out stealing or picked off, advances on a wild pitch or passed ball, etc.)?

Some argument could be made that the batter does contribute in some way to the changed situation (e.g., taking pitches to allow a steal, missing on a hit-and-run leading to a caught stealing, etc.). However, these changes generally have less to do with the batter, and more with the runners and the defensive team. As such, resetting the possible base advancement figure for the at bat probably makes the most sense.

Likewise, should a hitter be penalized when a runner is thrown out trying to take an "extra" base (e.g., going from first to third on a single)?

It would not seem fair for a batter who singles with a runner on first to end up with a 0 base advancement figure (+1 for the batter and -1 for the runner) if the runner is thrown out trying to go to third. It would seem to make the most sense to credit the batter for bases advanced by the runner prior to the runner being thrown out trying for the extra base, without subtracting any bases for the runner's out. (Obviously, force outs are not "trying for an extra base", and thus would not be included in this.)

A less clear case would be one where a runner is thrown out trying to advance one base in a nonforce situation. Some argument could be made that "counting" those bases lost due to the out would be justified, since the batter arguably contributed by presumably not hitting a ground ball behind the runner or hitting a deep enough fly ball when the situation called for it. However, the runner is probably a bigger contributor to the situation than the

batter. It might be best to simplify things by only subtracting bases lost due to runners' outs when the runner is out on a force out. (Note that doing this would also eliminate the penalty to the batter when runners are doubled off on line outs.)

(Note: One interesting twist to the BAP might be to allow some base running activities to flow into a player's BAP. A stolen base could add 1 more base advancement to the player's base advancement total, while a caught-stealing or a base running out while trying to advance an extra base could result in the loss of base advancements. However, this would cause a problem in the case of pinch runners, who could add to their base advancement totals without any increase to the potential base advancement totals that form the denominator of the statistic.)

Should any adjustment be made for at bats that result in a walk, hit by pitch, sacrifice, or sacrifice fly?

Both batting average and slugging percentage exclude these at bats from their calculation. However, the BAP would give at least some reward (advancement of at least 1 base) in each of these situations, so including them would make more sense in the BAP calculation. A possible exception might be an intentional walk, since that most often occurs when good hitters are at bat in clutch situations, and therefore might penalize that hitter even though giving some credit) by producing a low BAP for that at bat.

How should the result of errors be treated in the BAP calculation?

Both batting average and slugging percentage do not give credit to the batter when he reaches base due to an error. A batter does not get credit for an RBI when a runner scores as a result of an error. So it could well be argued that a hitter should not get credit for any

base advancements (by the batter or a runner) that are the result of an error. However, this could require official scorers to make additional judgment calls beyond those they currently have to make. For simplicity's sake, it might may be easier to let all base advancements flow into the BAP calculation.

In summary, the base advancement percentage measures how successful a player is in advancing himself and his teammates around the bases. In its simplest form, it would measure the player's success in doing this while batting. It could be expanded to include to some extent the player's base running actions as well. It may be superior in some ways to more commonly used offensive statistics, such as batting average and slugging percentage, in that it gives some reward for the little things, such as hitting behind the runner, and would seem to give greater weight to results in "clutch" situations.

Bill James' response:

OK . . . well, maybe you two guys can get together and you can negotiate your differences. Or maybe you should hire lawyers and sue each other, I don't know.

The (general) idea might be very worthwhile. In practice, the specific problems might sink it, or they might not; I really couldn't guess. The Achilles heel of win percentage systems is interactive effects—that is that, in practice, it is so difficult to figure out who exactly should be credited with what.

Suppose that there is a runner on first when a single is hit. A fast runner may go to third base; a slower runner may stop at second. A

stupid runner may be thrown out at third. An awkward runner may fall down before he reaches second.

This is but the beginning of the complication. If the runner from first DOES go to third, it is not at all unlikely that the outfielder would throw to third, allowing the batter to go to second. It is not terribly unlikely that the batter might try to go to second and get thrown out.

Thus we have a series of possible outcomes, on all of which the batter has done essentially the same thing—hit a single—and on all of which the batter would be “scored” as doing the same thing in a traditional accounting. The wide variety of possible scores, from 4/7 to 0/7, result from a context which is actually very different being accounted for as if it was always the same.

Whether this would be a practical problem, or whether it is merely an annoying theoretical obstacle . . . there is just no way of knowing. Even when you had the data, you would have to analyze it and debate it for years before you would reach any consensus on this issue.

I probably shouldn't tell you this, but for the Red Sox last year I did a vaguely similar analysis of base running, focusing on whether base runners picked up available bases—how often they moved up on wild pitches, how often they moved up on ground outs, etc. We discovered that one of the slower players on our team (Doug Mirabelli) was one of the most effective base runners on our team, at least by this analysis. But is this a real phenomenon, or is it a data fluke? Maybe they just happened to throw several wild pitches when he was on base? Maybe there were just an unusual number of situations when the single was hit in such a manner that he could make it to third base? Is it real, or is it a statistical illusion?

We don't know. We all watch the Red Sox, every game, veteran baseball men . . . we just don't know. Unless and until we have years worth of data, we won't know.

The analysis you are proposing is a lot more complicated than that—so it will take you MORE years and MORE data before you really know what you've got.

Then there is this question: is measuring a player against the potential inherent in a situation actually a valid method, or would it, in practice, be vulnerable to some bias you haven't anticipated?

Much of the analysis that we do is based on static changes—this state exists BEFORE the play (X1), this state exists AFTER the play (X2), therefore the batter's contribution is $X2 - X1$. That analysis has the same sorts of problems I have been discussing here—the constant difficulty in determining, in practice, who should be credited with creating the change of states.

It seems to me, intuitively, that when you turn the analysis around and measure it by POTENTIAL bases to be gained, rather than actual bases gained, you might very probably be introducing a catastrophic bias. But unless and until the work is actually done, there is just no way of knowing.