

Correlation for Better Courses

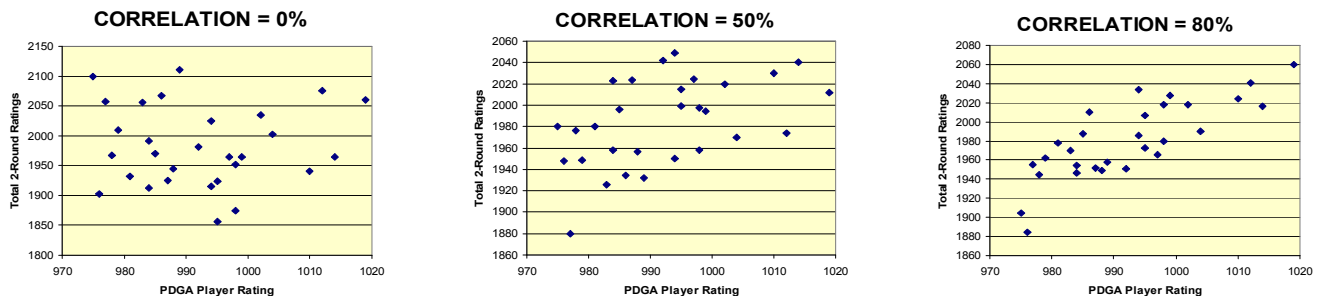
by Ratings Team

Can we come up with a number that determines how good a course is for competition, especially Majors? Courses are currently selected for our Majors based on a variety of subjective and objective criteria. Much of the time, the course(s) are simply the ones judged toughest in the city hosting. But does that automatically make them the best courses for determining a champion? *Is there a numerical value that can be calculated to objectively measure the "goodness" of the courses used to determine winners not only in Open but for each division regardless of skill level?*

Before diving into some calculations, let's assume the top 75 players entered in Open have established ratings where we can rank them in order at the start from top to bottom. After playing the event, they'll have a new ranking based on their total scores. In a perfect world on perfect courses with robot players, their finish order might always be the same as their starting order. Fortunately, it doesn't work that way in the real world. Otherwise there would be no reason to compete with the outcome always known in advance.

Now, if the finish order ends up close to the reverse order of their initial ranking (lowest rated player on top and highest at the bottom), we suspect something must really be out-of-whack with the courses, rules and/or format. Even in the case where players finished in random positions, as if they drew their finish position from 1 to 75 out of a hat, it would seem highly suspicious. Luck would appear to be the primary factor in determining their finish position with apparently little to no skill involved.

So what kind of correlation would seem acceptable, good or great between the players' starting rank and their finish rank that would make you think the competition was fair and based more on skill than luck? The stats world can help us with a calculation, not surprisingly called a *correlation*, that produces a number ranging from -1.0 to 0 to +1.0 to indicate how well two ranked lists correlate with each other. The robot example would be a perfect correlation of +1.0, the reverse order example would produce a perfect negative correlation of -1.0 and the random scatter finish where your finish position seemed drawn from a hat would produce a correlation near 0.



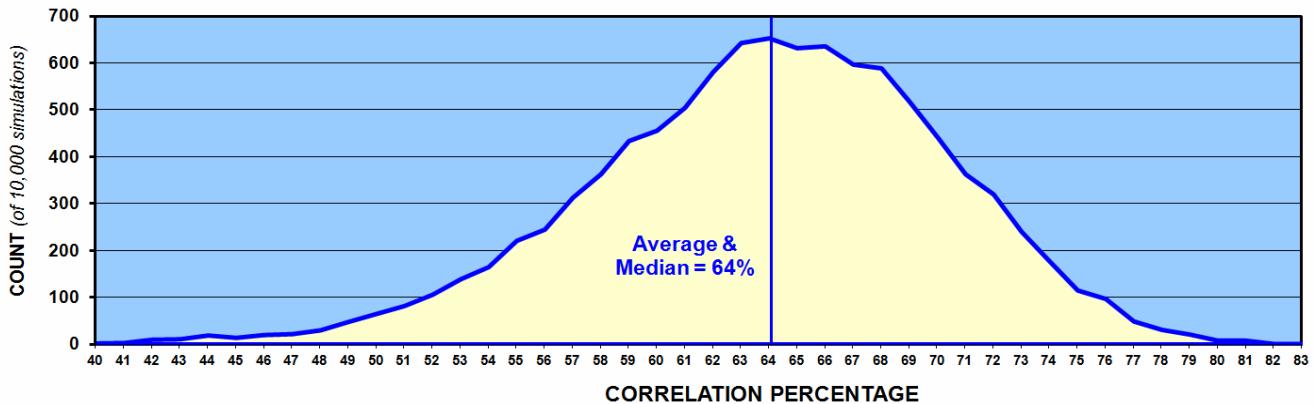
We'll use percentages instead of decimals to represent correlation values as we move deeper into this stats discussion. The graphs above show what 0%, 50% and 80% correlations might look like between a player's total rating for 2 rounds and their player rating. The 0% looks like a random buckshot pattern. As we move toward 80% correlation you can see the higher your player rating the more likely your 2-round ratings will be higher. We're going to see whether there's a correlation value as we approach 100% that tells us a course did a good job for a specific rating range of competitors.

Despite nicknames like Borg, we know players cannot always consistently play like robots. However, we have discovered that players with established ratings shoot scores in a range that can be predicted over 10 rounds even if predicting their exact score in their next round is a mystery.

The ratings team maintains stats on this probable scoring range for every player (called a standard deviation.) The team has produced a model that can randomly produce probable scores for up to 75 players based on their ratings and standard deviations for an event from 1 to 6 rounds. From this, we can produce a correlation distribution of their final ranks for 1000 event simulations at a time.

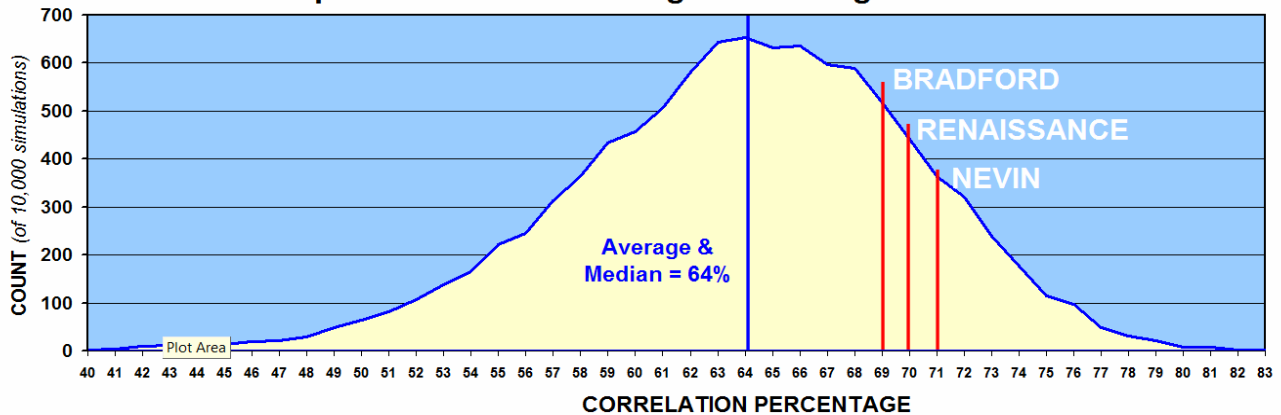
The graph below shows the expected correlation distribution for the top 75 Open players at Charlotte Worlds from 10,000 simulations producing total ratings for 2 rounds. Both the average and median correlations shown as the vertical blue bar come out about 64% with a range of +/-15% for a 2-round event on an "average" course. In other words, if every player played a hypothetical course equivalent to the ones that produced their rating, we would expect their finish rank to correlate with their starting rank to fall somewhere between 45% and 80% with 64% about average.

**Charlotte Worlds 2012 - MPO Expected Correlation
Player Ratings with 2-Round Ratings on "Average Course"**



Now let's compare actual 2-round ratings correlations from the three different courses played by this group of Open players to their expected baseline correlation range (yellow shaded area) and see where they fit. The correlations for the 2 rounds MPO played on Bradford, Renaissance and Nevin before semis are shown as the red vertical lines on the graph below.

**Charlotte Worlds 2012 - MPO 2-round Rating Totals
Expected Correlation Range for Average Course**



These single correlation data points calculated for the three courses look very good for starters at 69, 70 and 71 percent correlation. Those are all better than our hypothetical average course correlation average and also stand alone as good correlations indicating each course played well. However, one data point for a course doesn't have enough predictive value to necessarily indicate if these courses would always play well for MPO. But it's a start.

Ideally, we would want to compile MPO scoring data from several 2-round events on each course to get a higher level of confidence that their cumulative correlation percentages were truly better than the "average" course for these players. But for now, let's take a look at some single correlation data points for courses played by Open in 11 of the past 14 years at Pro Worlds. Pro Worlds in 2000 and 2011 were skipped since their 24 and 27 hole data isn't comparable with the 18-hole course data from the other 11. Some courses were also not included if they were not played twice by Open before semis.

The first column (**Correl**) in the following table ranks the correlation in descending order between the player ratings for up to the top 75 Open players in those World Championships with their total ratings after playing each of these courses twice. Highbridge Gold at 76% had the highest correlation of any Worlds course played by Open since ratings began in 1999.

Correlation of Total Ratings for 2 Rounds with Open Player Ratings

Correl	Course	Year	SSA
76%	Highbridge Gold	2007	65
73%	Lemon Lake Gold	2010	63
71%	N. Arizona U	2003	51
71%	Bradford*	2012	53
70%	Renaissance	2012	65
69%	Nevin	2012	58
69%	Timber Ridge	2008	62
69%	Victory	2008	52
66%	North Valley*	2001	54
66%	Lake Olmstead	2006	54
66%	Blueberry Hill	2007	54
65%	Oakwood	2001	55
65%	Cliff Drive	2009	55
63%	Little America	2003	53
63%	Red Hawk	2010	53
62%	Genessee Valley	1999	48
62%	Black Creek	1999	48
62%	Pickard*	2004	56
60%	Oshtemo*	2008	55
60%	Nockamixon	2005	67
60%	Granite Ridge*	2007	57
59%	Jordan Creek	2005	56
58%	Water Works*	2009	52
57%	Hippodrome	2006	63
56%	Ellison	1999	49
56%	Coldbrook	2008	50
56%	Little Lehigh*	2005	56
55%	Blue Heron*	2010	52
54%	Kaposia	2001	54
54%	Tinicum	2005	55
53%	C.P. Adams	2001	54
52%	Riverview*	2006	51
51%	Chili*	1999	50
45%	Iowa State	2004	49

* Semi-Final course

Course	Pct.
Highbridge Gold	95%
Genessee Valley	92%
Bradford*	91%
Black Creek	90%
Renaissance	86%
Nevin	81%
North Valley*	79%
N. Arizona U	76%
Oakwood	71%
Ellison	70%
Timber Ridge	60%
Victory	60%
Lemon Lake Gold	55%
Chili*	50%
Pickard*	35%
Blueberry Hill	34%
Oshtemo*	24%
Little America	23%
Kaposia	15%
Nockamixon	12%
C.P. Adams	12%
Lake Olmstead	12%
Water Works*	11%
Jordan Creek	10%
Granite Ridge*	10%
Cliff Drive	7%
Red Hawk	3.7%
Coldbrook	2.8%
Tinicum	2.6%
Little Lehigh*	2.4%
Riverview*	0.6%
Iowa State	0.4%
Hippodrome	0.3%
Blue Heron*	0.1%

The columns on the right of the previous table show the ranking for how well each 2-round course correlation was in comparison to the correlation range for the average course for this pool of players. For example, the Bradford correlation was better than 91% of the expected average course correlation range for the top 75 PW2012 Open players.

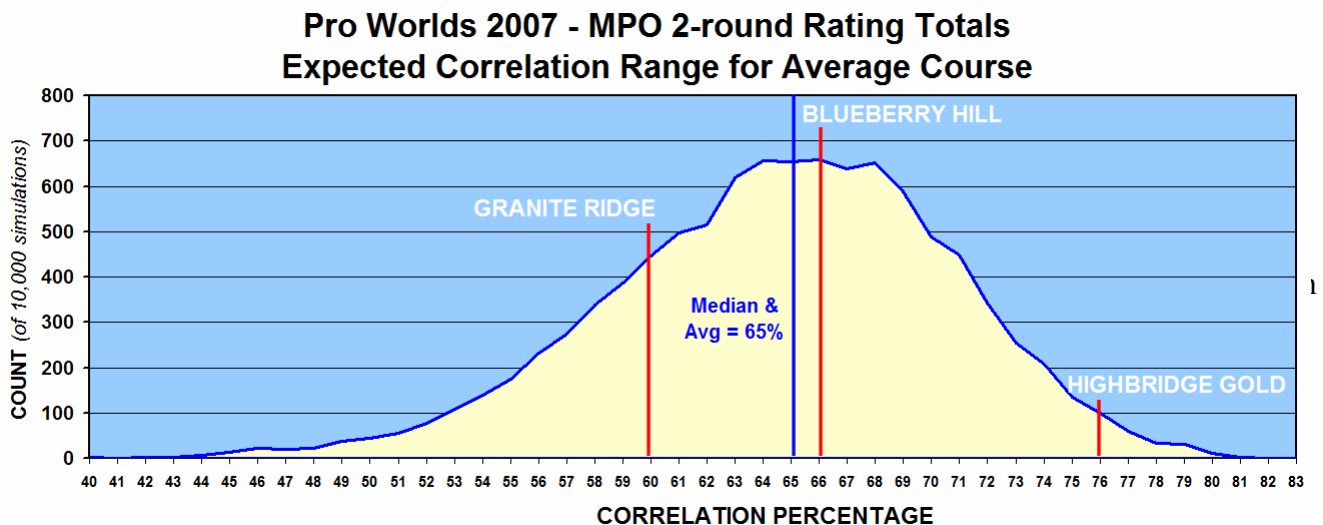
Player ratings were less established in the first few years until 2003. So it may be the reason the correlation range for the average course ended up a bit lower for the 1999 and 2001 courses.

Do we see any characteristics of the courses that seem to produce better correlation values? That will take much more data to determine since again we currently only have one correlation data point per course. So far, there's a mild correlation of 45% between SSA (Scratch Scoring Average) and the correlation value produced. This might be expected since longer and tougher courses in theory should produce a wider range of scores to separate players by skill.

For example, four of the six courses with SSAs over 60 had correlations in the top seven. The two with lower correlations are likely the most tightly wooded of the six. Except for 2006, the course with the highest SSA at each Worlds had the highest correlation. At the other end of the spectrum, it may be no surprise to some that the Iowa State course with a 49 SSA ended up as the weakest for Open. It was a short, new course torched with a 41 & 42 averaging 1080 by Barry Schultz. The hosts didn't have a chance to test and possibly tweak any hole designs before Iowa Worlds. Another problem with temp or new courses being used in competition is that off-fairway "shank" areas can be nasty with no time for traffic to wear them down, thus reducing chances to execute exceptional recovery throws.

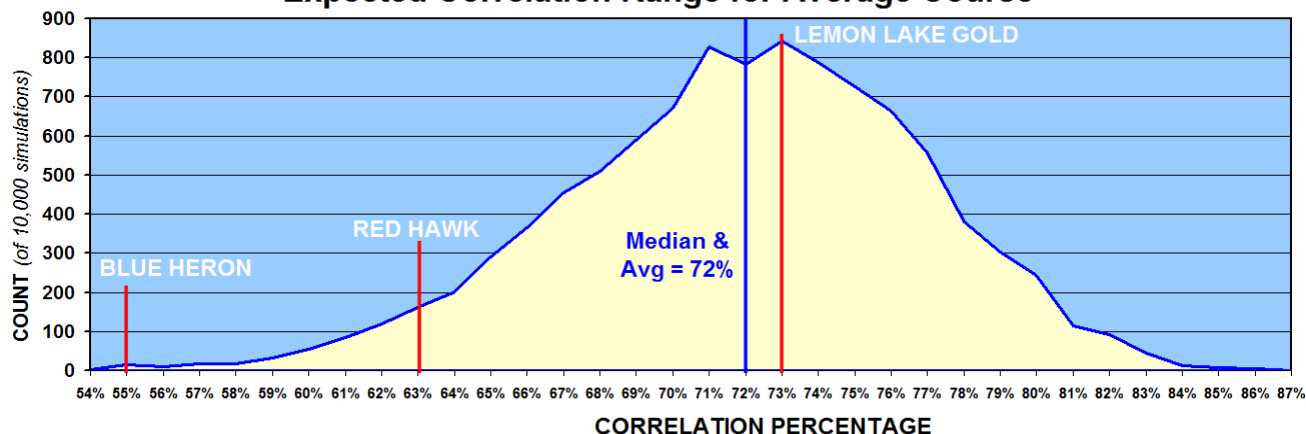
It's interesting to look at the courses used for Semifinals which are indicated with an * asterisk in the table. Only half of the Semifinal courses had a correlation at least 60%. As a group those had higher SSAs than the lower ranked ones. It's normal policy at Worlds to either play an "easier" course or shorten them for Semifinals so the rounds go faster and players can make it to the Finals. However, this policy may mean the course chosen for semifinals may not always be the best at least for MPO.

The **Pct.** column in the right side of the table is the percentage of area under the average course distribution curve that's to the left of the correlation mark for that course. The following graphs may help visualize where these courses fit compared to their "average course" distribution ranges. You can see that most of the yellow area (95% of it) is to the left of the Highbridge Gold red line.



Note that the median and average values at 65% for this PW2007 pool of Open players are about the same as the 64% for the Charlotte MPO players in the graph on page 2 even though their average rating in 2012 was 1011 versus 1002 in 2007. However, the following graph for Pro Worlds 2010 shows the average course correlation range can shift up or down and isn't automatically set in the 64-65% range each year. In 2010, the average and median moved up to 72% with a pool of Open players averaging 1004 rating, halfway between the 2007 and 2012 Open pool average ratings. What it could mean is the Open pool in 2010 had more players with smaller standard deviations (more consistent). That raised the bar for courses to beat the average expectation for correlation that year.

Pro Worlds 2010 - MPO 2-round Rating Totals Expected Correlation Range for Average Course



The Red Hawk course had a decent 63% absolute correlation but still had a low **Pct.** value due to the higher overall correlation expectations for this field of MPO. Blue Heron barely made it on this average course distribution graph. It's still just one data point in terms of predicting whether a course is suitable for Open in the future. But it can definitely be a cautionary flag until more data is gathered. The Lemon Lake complex will again be hosting Pro Worlds in 2013. The host team is taking a look at course improvements and using two other possibly hybrid courses for Open, instead of Red Hawk and Blue Heron, in addition to Lemon Lake Gold partly due to feedback from this analysis.

Preliminary Round Correlations at Worlds

As more rounds are played in an event, the correlation percentage will progress higher because players find it increasingly difficult to escape averaging closer to their player rating. The following table shows the correlations for 6 preliminary rounds (or the equivalent) played by up to 75 top Open players at Worlds. Even the lowest 6-round correlation at 73% is almost better than the highest 2-round correlation of 76% in the previous table.

There's little correlation (6%) between the average MPO SSA and the 6-round correlation for top 75. Another curiosity is that having good 2-round correlations for the individual courses does not necessarily produce a higher 6-round correlation than other Worlds that had lower average 2-round correlations for their courses.

Worlds Location	Worlds Year	MPO Average SSA	6 Rounds Equiv. Correlation	
			Top 75	All MPO
Charlotte, NC	2012	58.9	82%	91%
Kansas City, MO	2009	56.4	82%	87%
Houston, TX	2002	51.6	81%	85%
Flagstaff, AZ	2003	52.0	77%	85%
Santa Cruz, CA*	2011	56.3*	77%	79%
Allentown, PA	2005	58.9	76%	85%
Kalamazoo, MI	2008	54.6	76%	85%
Crown Point, IN	2010	55.1	75%	87%
Augusta, GA	2006	54.8	75%	87%
Twin Cities, MN	2001	54.0	75%	79%
Highbridge, WI	2007	58.5	73%	87%
Des Moines, IA	2004	53.5	73%	82%

* 18-hole equivalent for four 27-hole courses

For example, the three 2007 courses had 2-round correlations higher as a group than the four 2005 courses. And yet, the 6 round equivalent correlation in 2005 was a few points higher than 2007, 76% to 73%. On the other hand, the high 2-round correlations for Charlotte courses did produce the highest 6-round correlation. The relative similarity of courses in Charlotte may be the reason. A more diverse set of courses may produce a lower 6-round correlation because a slightly different set of players shoot well on each course that suits their strongest skill set.

For example, if all three courses are similar, let's say wooded courses, it's possible the same players are shooting well on each course producing good 2-round correlations. This would naturally produce a higher 6-round correlation with better players on wooded courses separating from the others. However, what if the three courses are markedly different, say one ski hill, one mostly open and one shorter wooded? It's possible a slightly different set of players could produce a good 2-round correlation on each course matching their slightly different skill sets. When these good correlations for each course are combined for 6 rounds, the scores end up tighter together. So the composite correlation doesn't increase as much.

The better 6-round correlations must be checked to see whether course similarity, even if each is good on its own, are so similar that they aren't fairly balanced as a group. A lower 6-round correlation, especially when each course is good on its own, may actually be a better balanced set of courses considering that the competing players' ratings were produced on all types of courses. More study will be needed to determine whether the 6-round correlation is meaningful on its own without further analysis.

Notice in the far column of the table that the 6-round correlation is even higher when you use the whole pool of MPO players rather than just the top 75. This is to be expected since the farther players are rated below the top players, their chances to win diminish to virtually zero. That's why a smaller group comprising up to the top 75 MPO players were used in this study. In fact, a case could be made that only players likely to be in contention or who are likely to cash should be included when this type of correlation analysis is done.

USDGC Analysis

The USDGC typically has the highest rated field of any event in a year. Even though it is always held on the Winthrop Gold course, the rules and some holes change each year. Thus, it hasn't provided a consistent format to compile correlation results that can be combined over multiple events. We can only look at single correlation data points for each year.

The following table shows some correlation values from five of the most recent six USDCG events, skipping the 2011 Performance version. What we discovered in the process of analyzing these numbers is the range of player ratings included in the correlation should be set close to the same number each year. It was set close to a 40-point range for this analysis with no player under 1000 rating included.

Year	Rules Characteristics	Correlations			Pct.	Player Rating			Player Count
		1st 2-rnds	2nd 2-rnds	Avg 2-rnds		Hi	Lo	Range	
2012	Distance Only OB (Buncr)	55%	55%	55%	54%	1046	1005	41	41
2010	All Throw & Distance OB	62%	57%	60%	92%	1043	1002	41	57
2009	Buncr + Throw & Dist. OB	41%	54%	48%	54%	1041	1005	36	57
2008	Buncr + Throw & Dist. OB	50%	65%	58%	91%	1039	1001	38	55
2007	Multiple Island Greens	30%	53%	42%	21%	1039	1000	39	51

With the USDGC including 4 rounds, at least for the top players each year, there were two sets of 2-round scores to produce the correlations as shown in the table. Other than 2012, the correlation values were significantly different between the first 2 rounds and second 2 rounds and not always better. This variance is what we would expect regardless of how good or weak a course may be for a player group. It's only after accumulating several data points that there's enough data to see a more precise average correlation emerge.

Even though the 2-round correlation percentages are middle-of-the-road compared with the Worlds range on a previous table, some years were pretty good in comparison to the expected range for the smaller USDGC field selected and ratings range compared to the Worlds values. For example in 2008 and 2010, more than 90% of the expected correlation range (see **Pct.** column) for that field of players fell below the average correlation values. This indicates the course rules and layouts worked pretty well those two years to challenge and rank the better players.

Where do we go from here?

Correlation analysis is a tool the PDGA could apply to evaluate much of our old and future tournament database to determine how well courses rank players of all skill levels in tournament play. This process isn't just useful for the top players but every player. Ideally, we would want to discover that our Gold level courses have good correlations when top players play them and our White level courses have good correlations when Intermediate and Rec players play them.

The ultimate benefit from all this number crunching could be realized when we figure out which course characteristics seem to produce good correlations and which ones lower them. That will take an in-depth effort over the long haul. But the big payoff would be providing guidance to help improve our courses for all players.