Ψ: An Alternative Measure of Disc Golf Player Performance

by Steve West Disc Golf, LLC

December 24, 2020

The goal is to come up with a value which could represent how a player performed over a set period of time. We'll call it Proportionate Score Index, or P.S.I. for short, or Ψ for shorter. The value of Ψ would be based solely on head-to-head play between pairs of players.

Ψ should tell us exactly how well each player actually played against the field of players they faced, and should give us a good idea of how well they would play against any single player whether they have met or not.

Another desirable aspect of Ψ would be that every throw by a player counts equally in contributing to the value of Ψ.

Dealing with Data Load

Most of the difficulty of such a measure is handling the huge amount of potential data. I have data for 4,365 players from 2020. That's over 16 million potential combinations of head-to-head competition. Add in the fact that players may meet in several tournaments and the number of cells it would take to keep track of it all is staggering.

Following are some ways to whittle that down.

To start, we don't need to set up a grid for all the potential matchups, we just need to track those that actually occurred.

We don't need hole-by-hole scores, either. Total score for the round will tell us all we need to know. Extending that thought, if both players played multiple rounds on the same layout, all we need is the scores from all of their rounds on that layout during that event. Extending that thought further, for any two players all we need to know is the sum of all scores they got on all the rounds they played in the same event on the same layouts.

(A complication arises when the two players did not play the same number of rounds on a layout. To solve that, I used the average score for each player. For purposes of weighting the data for credibility, I used half of the total number of rounds played by both players.)

Because identifying unique players was crucial, any players without a PDGA Number were excluded. However, they did not need a PDGA Rating to be included, because Ψ is based only on scores.

Some events were not included due to complications in figuring out who played what layout when. Also, I did not collect much data for Amateur players.

The average player in the database made 334.6 throws; during 5.2 rounds; for an average score of 64.0; on 3.2 layouts; and had an average rating of 926.7 over the year.

The data includes 276 different event/course/layouts. There were 13,763 different combinations of player+layout. Each player's scores on a certain layout will match with a number of other players' scores on that layout. All we care about is how they performed head-to-head, so all the scores for any set of two players on all their shared layouts can be summed. This brings the number of head-to-head matchups down from the 19 million possible combinations to the 389,351 which summarize what actually happened.

(The data could easily be expanded as more events are played by continuing to add on to the total scores of any two players that meet on the same layout.)

Sample Head-to-Head

As an example, Paul McBeth competed with 600 unique players.

Here are the top ten who came closest to tieing with Paul.

Player A	Score A	Player B	Score B	Ratio
Paul McBeth	1888	Calvin Heimburg	1897	99.5%
Paul McBeth	1726	Richard Wysocki	1741	99.1%
Paul McBeth	1499	Eagle Wynne McMahon	1521	98.6%
Paul McBeth	1587	Chris Dickerson	1613	98.4%
Paul McBeth	191	Øyvind Jarnes	198	96.5%
Paul McBeth	426	Michael Johansen	442	96.4%
Paul McBeth	1888	Kevin Jones	1965	96.1%
Paul McBeth	165	Andy Martin	172	95.9%
Paul McBeth	1717	Garrett Gurthie	1791	95.9%
Paul McBeth	528	Cale Leiviska	551	95.8%

These 8 played the most layouts with Paul.

Player A	Score A	Player B	Score B	Ratio
Paul McBeth	1888	Calvin Heimburg	1897	99.5%
Paul McBeth	1888	Kevin Jones	1965	96.1%
Paul McBeth	1888	Paul Ulibarri	1994	94.7%
Paul McBeth	1888	Austin Hannum	2005	94.2%
Paul McBeth	1888	Jeremy Koling	2024	93.3%
Paul McBeth	1888	Eric Oakley	2035	92.8%
Paul McBeth	1888	Ezra Aderhold	2058	91.7%
Paul McBeth	1888	Terry Rothlisberger	2112	89.4%

Where Paul's competitors threw 249,020 throws, Paul threw only 215,114 or 86.4% as much. However, Paul faced a higher level of competition than most players, so he is really better than the 86.4% would indicate.

Solving for Ψ

When $Player_1$ plays $Player_2$, the expected ratio of $Player_1$'s score to $Player_2$'s score is (Ψ_1 / Ψ_2). To solve for Ψn, every player starts with Ψn=1. Then, the resulting expected score for $Player_1$ is calculated based on the total throws made in each matchup times $(\Psi_1 / (\Psi_1 + \Psi_2))$. The value of Ψ_1 is adjusted until $Player_1$'s expected throws equals $Player_1$'s actual throws. Ψ_2 also gets automatically adjusted in this process, because all players are $Player_1$ half the time and $Player_2$ half the time.

The resulting values for Ψ average out to 1, with a range of about 0.8 to 1.3. We can translate these into a more relatable value by multiplying Ψ times the average score of all players in the data set. Call this the Relatable Score.

Here are some values for players with at least 10 rounds in the data:

The top ten players, with their Ψ and relatable score are:

Player	PSI	Relatable Score
Paul McBeth	0.809	51.8
Richard Wysocki	0.817	52.3
Eagle McMahon	0.819	52.4
Calvin Heimburg	0.822	52.6
Chris Dickerson	0.829	53.1
Väinö Mäkelä	0.837	53.6
Nathan Sexton	0.841	53.8
Andrew Fish	0.843	54.0
Kevin Jones	0.847	54.2
Drew Gibson	0.850	54.4

The top ten "Big Fish in a Small Pond" (Ψs lower than the field) are:

Name	PSI	Relatable Score	Opponents'	Difference
Richard Wysocki	0.817	52.3	61.1	-8.8
Calvin Heimburg	0.822	52.6	60.9	-8.3
Paul McBeth	0.809	51.8	59.9	-8.1
Chris Dickerson	0.829	53.1	61.1	-8.0
Niklas Anttila	0.853	54.6	62.5	-7.9
Eagle McMahon	0.819	52.4	59.5	-7.1
Kevin Jones	0.847	54.2	61.4	-7.2
Jesse Nieminen	0.864	55.3	62.5	-7.2
Albert Tamm	0.863	55.2	62.4	-7.2
Cale Leiviska	0.86	55	62.1	-7.1

The top (or bottom?) players in over their heads (Ψs higher than the field) are:

Name	PSI	Relatable Score	Opponents'	Difference
Alyssa Pierson	1.234	79.0	61.7	+17.3
Christine Huestis	1.298	83.1	66.7	+16.4
Lindsey Langley	1.2	76.8	62.2	+14.6
Kayla Barron	1.234	79.0	64.7	+14.3
Joe Bishop	1.145	73.3	59.9	+13.4
Chelsea Harden	1.169	74.8	61.9	+12.9
Margaret Baudendistel	1.171	75.0	62.7	+12.3
Lindsay Fish	1.149	73.5	62.0	+11.5
Thomas Cupp	1.089	69.7	58.5	+11.2
Madison Tomaino	1.13	72.3	61.6	+10.7

While having nothing to do with Ψ, the hardest-working players (those who threw the most throws during the competitions included in the data) were:

Name	Total Throws
Catrina Allen	3,965
Cameron Messerschmidt	3,816
Holly Finley	3,777
Austin Hannum	3,716
Emerson Keith	3,645
AJ Carey	3,360
Terry Rothlisberger	3,310
Missy Gannon	3,220
Ezra Aderhold	3,212
Kona Star Panis	3,180

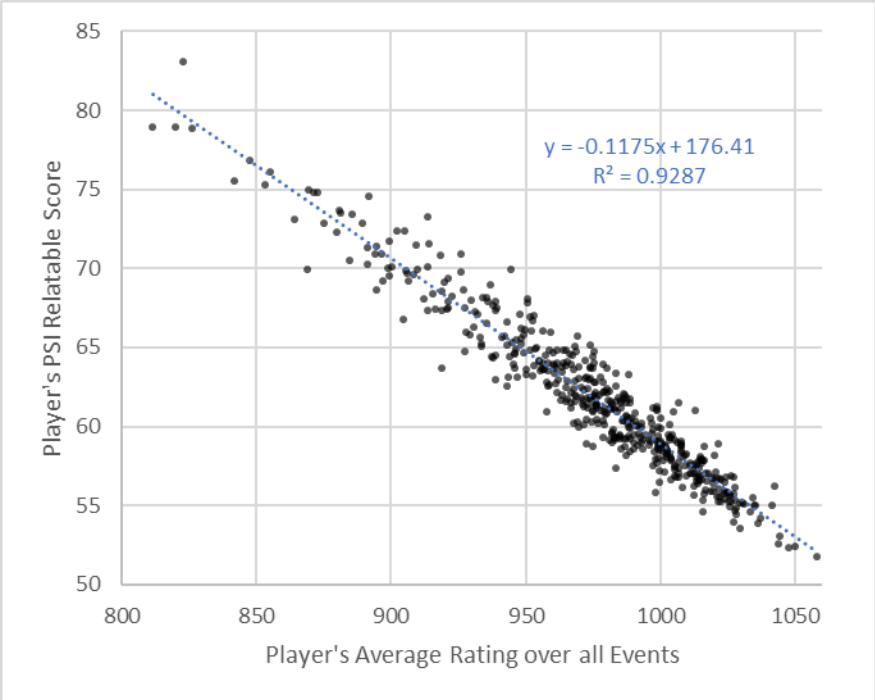
The top ten Cherry-Pickers (players who played on courses where they got low scores in relation to their Ψ) were:

Name	Relatable Score	Actual Average Score	Difference
Justin Scoggins	69.9	59.7	-10.2
Jerry Goff	63.7	53.7	-10.0
Kayla Barron	79	70.3	-8.7
Nathan Lavender	65.2	56.8	-8.4
Bamba Rico	64.8	56.5	-8.3
Brandie Myers	73.4	65.6	-7.8
Scott Withers	56.2	48.5	-7.7
Jennifer Allen	67	59.3	-7.7
John Kotansky	61.6	54.3	-7.3
Joe Carey	63.1	55.9	-7.2

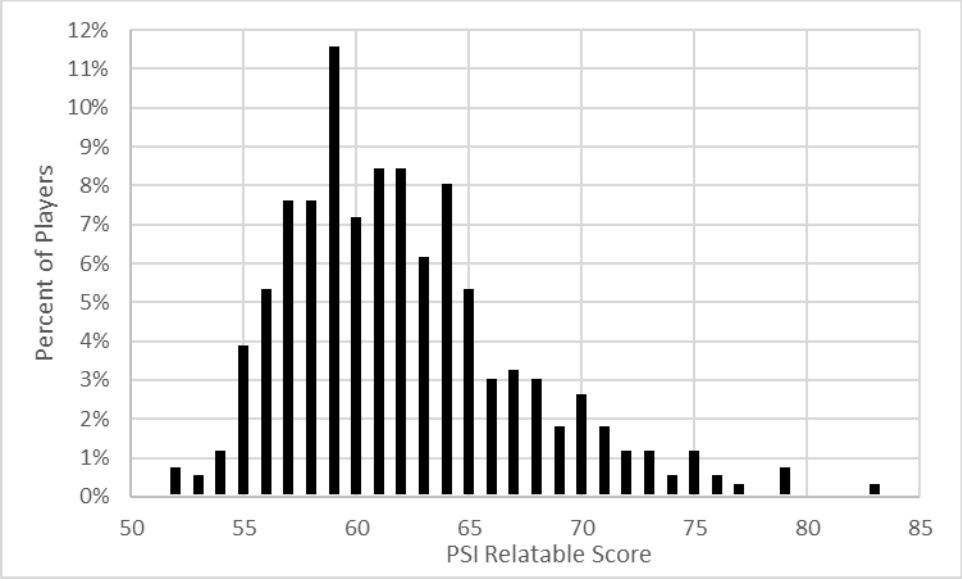
At the other end, these players played on courses where their scores were highest relative to their Ψ (in the form of Relatable Score).

Name	Relatable Score	Actual Average Score	Difference
Travis Tschida	65.6	77.9	12.3
Logan Utter	67.4	77.0	9.6
Jesse Adams	60.2	68.3	8.1
Jared Johnson	59.8	67.5	7.7
Jeremiah Dwyer	61.8	69.0	7.2
Ryan Muizelaar	58.0	64.9	6.9
Grant Dammann	63.3	70.2	6.9
Will Bratzel	62.0	68.7	6.7
Gavin Babcock	58.8	65.3	6.5
JohnE McCray	55.6	61.1	5.5

The chart below shows how Ψ (in the form of Relatable Score) compares to each player's average PDGA Player Rating for all the events that went into the data (for those players with a rating). As expected, the two measures of skill are largely in agreement. There are two main differences. First, Ψ puts the same weight on every throw. Second, PDGA Player Rating includes past performance, not just performance during the period under study.



Here is the distribution of skills measured by Ψ (in the form of Relatable Score):



Predictions

Ψ can give the expected ratio of the total throws. Thus, Ψ can predict who would win in a head-to-head match: for a hole, for a round, for an event, or for a season.

The average PDGA Player Rating cannot be readily translated into an expected scoring differential. The only thing it can tell us is who should win (the player with the higher average PDGA Player Rating).

So, to compare the two, I computed how often the winner of a season-long matchup had a higher PDGA Player Rating, vs. how often the winner had the lower Ψ (or Relatable Score).

I only looked at matchups where both players had a rating and the two players did not tie.

For matches where there was a winner, average PDGA Player Rating predicted the winner 84.4% of the time, while Ψ predicted the winner 93.8% of the time.

Conclusion

The arithmetic is simple, the concept is easy to understand (when expressed as a Relatable Score), the resulting measure of performance is better than PDGA Player Rating, the data is easily available, and many spin-off interesting nuggets can be generated. The main difficulty is handling the sheer volume of data.