RJR Ratings System
Executive Summary

World English-Language Scrabble® Players Association (WESPA)

Version 4.0

Russell Honeybun

Yong Jian Rong

8 October 2024

Introduction to the RJR Ratings System

1. What is the RJR Ratings System?

"RJR" refers to the primary authors, **R**ussell Honeybun and **J**ian **R**ong Yong. The RJR rating system is based on the Glicko rating system, featuring finely tuned variables that have been refined specifically for the WESPA Scrabble-playing population. While the authors refer to their own system as “Glicko k=250 winspace1 v4”, on advice of the WESPA Chair we will refer to it as the RJR system for simplicity.

This new system is a culmination of research and multiple trials done since the start of 2024 and was approved by the WESPA Committee at its July 2024 monthly meeting. Implementation will occur on 1st January 2025.

2. Why has the rating system been changed?

A review of the incumbent ratings system was conducted to investigate and address perceived issues related to youth player performance. Population and time-analysis performed showed that the incumbent rating system was not adequately rewarding effort, and that many players were penalised for either playing as a youth early in their scrabble career, or for playing those youths who are now underrated. The infrequency of WESPA-rated tournaments makes it exceedingly difficult to regain those ratings in a reasonable timeframe.

How did we come up with the new system?

First, the authors discussed what they believed to be the hallmarks of a good ratings system. This included i) A rating system should be a "data" system. That means we can rely on it to inform us of facts. We should be able to play anyone knowing our chances of winning ii) , the system needs to be robust with reproducible methodology and results. iii) The system should reward performance and consistent effort iv) the system should not produce an unfair advantage to infrequent play or volatile results v) the system should be fair to the most number of players possible, regardless of geographical location vi) the average rating of the system should not change over time and vii) players should be able to compare their rating in 5, 10 or 50 years

Second, the authors designed and tested many different systems based on historical results. They chose the best one through rigorous, repeatable testing with the goal to minimise the difference between expected outcomes and actual outcomes, eventually choosing a system that produced the lowest error margins.

3. What is the new system?

The new Glicko system is based on publications by Mark Glickman, a senior statistics lecturer at Harvard University with over 40 publications in peer-reviewed journals. The system was originally designed for chess but has been used in many competitive sports and e-sports alike to grade performance. The authors adjusted the base model outlined specifically attuned to Scrabble competitions. More than 400,000 WESPA-rated games across 14 years of competition between 2006-2019 were used to model the final parameters.

4. How does the new system compare to the incumbent system?

Every effort has been made to make the new system as familiar as possible, the mechanics 'under the hood' are vastly different. The biggest change by far is that you now have two numbers that are important, your rating and your rating deviation (RD). Instead of thinking of your rating as a singular number, these numbers indicate a range where the authors are extremely confident your true rating will lie. For example, an established player with 1400 rating and RD 100 we expect that the player would have a performance between 1200-1600 in 95% of their tournaments.

When you play in tournaments, your rating and RD, as well as your opponent's rating and RD are equally important in the calculation of your expected wins and rating change.

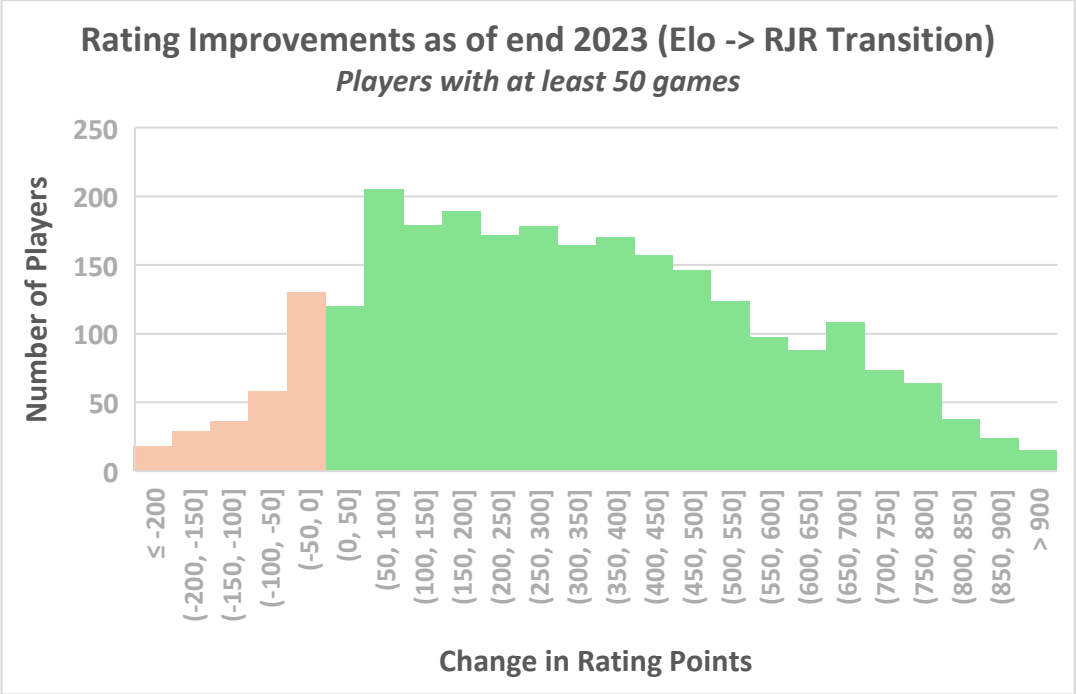
In addition to the above, not all players will have the same multiplier when applied to their rating gain. This allows newer players to be able to 'catch up' with the pack more easily, without damaging the rating of the established players. As a result, it is common for tournaments to result in net positive rating gain.

Your RD is affected by two things: how long it was since your last tournament, and your overall tournament performance. A lower number indicates a more reliable and consistent player. The highest ranked players in the world consistently achieve RD 50, which is a level the authors believe to be the perfect minimum value.

5. Changes to Player Rating Population

Out of 2583 players with ≥50 WESPA-rated games as of 31 Dec 2023:

- 2318 players will gain rating points, mostly ranging from +50 to +400.
- 265 players will have their ratings adjusted down to the new Glicko rating.



6. Impact on Win Expectations

Without going deep into the mathematical calculations, here are the win expectations (the chance that you are going to defeat an opponent) based on the rating difference between you and your opponent. Loss expectation can be calculated using '1 – win expectation'.

This is only a linear approximation that is accurate to 1% for each opponent. Refer to the next page for sample calculations.

Table 1 Win Expectation according to Rating Difference

Rating Difference	**Win Expectation**	**Change per 10-point increase**	**Rule of thumb for estimation** *(Linear approximation)*
10	0.5098	0.98%	10% per 100 points
20	0.5195	0.98%	
30	0.5293	0.97%	
40	0.5390	0.97%	
50	0.5487	0.97%	
60	0.5584	0.97%	
70	0.5680	0.96%	
80	0.5775	0.96%	
90	0.5870	0.95%	
100	0.5965	0.94%	
150	0.6425	0.92%	+9% per 100 points
200	0.6860	0.87%	
250	0.7265	0.81%	+7.5% per 100 points
300	0.7636	0.74%	
350	0.7970	0.67%	+6% per 100 points
400	0.8268	0.60%	

Note: This expectation table assumes an RD of 70 for both players.

7. Sample Calculation for Rating Change

Table 2 Rating Change Estimate per game won/lost

	Rating Change Multiplier per Game			
Pre-tournament RD	8-game tourney	16-game tourney	24-game tourney	32-game tourney
100	30	24	20.5	17.5
90	25.5	21	18	16
80	21	18	16	14
70	17*	15	13.5**	12
60	13	11.5	10.5	10
50	9	8.5	8	7.5
*see Example 1 **see Example 2				

These point values are rounded to nearest 0.5.

Win = 1, Draw = 0.5, Loss = 0

Rating change in a game = (Game Outcome - Expectation) x Rating Change Multiplier

The two examples below demonstrate the use of Table 2.

Example 1: Player A, rated 1700 and RD 70, joins a 8-game tournament

In a tournament where Player A meets 8 opponents of rating 1600 and their pre-tournament RD are 70, Player A has a win expectation of 0.5965[1] against each opponent, summing up to an overall expectation of 4.8 games.

If Player A wins 6 out of 8 games, estimated rating change = (6.0-4.8) x 17 = +20.4 points

If Player A wins 4 out of 8 games, estimated rating change = (4.0-4.8) x 17 = -13.6 points

Example 2: Player A, rated 1700 and RD 70, joins a 24-game tournament

In a tournament where Player A meets 24 opponents of rating 1600 and their pre-tournament RD are 70, Player A has a win expectation of 0.5965[2] against each opponent, summing up to an overall expectation of 14.3 games.

If Player A wins 18 games, estimated rating change = (18.0-14.3) x 13.5 = +50.0 points

If Player A wins 12 games, estimated rating change = (12.0-14.3) x 13.5 = -31.1 points

[1] Refer to Table 1 for win expectation for a 100-point rating difference.
[2] Refer to Table 1 for win expectation for a 100-point rating difference.