

# Ratings System Review

Implementation Guide

Prepared by: Steven Brown

October 2017

# RATINGS SYSTEMS REVIEW

## Background

The management meeting at the 2014 Nationals asked Glenda Foster to convene a panel to review the ratings system. Glenda advised the management meeting at the 2015 Nationals that she had nothing to report, and felt that I would be in a better position to review the system. Over the following year I was reviewing the system, posting material to the #ratings-review channel at nzscrabble.slack.com, and seeking feedback from others interested in the review. I presented an interim report to the management meeting at the 2016 Nationals, and produced a final report, incorporating feedback on the interim report, for consideration at the October 2016 management meeting (at the Mt Albert tournament).

## Final Report

The final report recommended:

- i) That the changes and transitional measures proposed in the final report be adopted for tournaments, starting with the first tournament in 2017.
- ii) That I produce a cross-platform implementation of the new system, that can be used to easily administer it.
- iii) That the new system be reviewed after two years of operation.

I have produced an implementation of the new system that closely follows the proposed changes to the new system. The few instances where the proposed changes were not able to be implemented exactly as proposed will be detailed below.

The implementation has currently been compiled to work under Mac OS X, and could be compiled to work under Windows or Linux without too much bother.

## Recommendations of Final Report

*Adopt NASPA/WESPA curve (a logistic curve with a scaling factor of 313).*

*Modify K factor from  $(3000 - \text{currentRating})/50$  to  $(3000 - \text{currentRating})/1000 * \text{ratedGames}$ .*

*Allow back-to-back one-day tournaments to be rated as if they are one two-day tournament.*

*Reduce provisional games to 30.*

*Allow all New, Provisional, Historic players to be re-graded (with agreement).*

*Rate Historic players on the same basis as established players.*

*Rate New and Provisional players using a combination of a performance rating from the current tournament and the previous provisional rating according to the ratio of the respective games played.*

*Add 'participation points' to each new rating (one point per three games rated).*

*Add 'accelerator' and 'feedback' points to the current system.*

Please refer to the final report for a description of the old system.

### Implementation of Proposed Changes

#### *Provisional ratings - ratings curve.*

I recommend that the number of games where the provisional rating methods are applied for tournaments where the player starts with less than that number of games be reduced to 30, as in the WESPA system. I also recommend that the provisional rating achieved after the tournament before the tournament in which a player passes 30 games be saved, along with the number of games in all prior tournaments, and be used to calculate the new provisional rating for each new tournament. The provisional rating from the tournament in which the player exceeds 30 tournament games then becomes their first established rating for entering their next tournament. The number of games included in the final provisional rating calculation will therefore vary between 30 (if the player starts an n-game tournament having played 30-n games beforehand) and 44 (if the player starts a 15-game tournament having played 29 games before).

I recommend changing the curve from the current Normal Standard Distribution curve to a logistic curve with the same scaling as used by both NASPA and WESPA. Provisional ratings would use this formula for each game being included in the provisional rating (including games against opponents new to the NZ system):

$$\text{OpponentRating} - \text{Log}(1 / \text{WinRate} - 1) * 313$$

where the OpponentRating is what the opponent in question's rating was at the time of the game, and the WinRate is for the tournament the game was part of. The resulting provisional ratings would be summed along with the total per game provisional ratings from previous tournaments and then divided by the total number of games. Log(p) is a widely available function giving a natural (i.e. base-e) logarithm. 313 is the empirically derived scaling factor arrived at independently by NASPA, WESPA, and my analysis of the NZ ratings database. The actual figure used by NASPA is close to 1/313, with multiplication and division swapped as appropriate in the formulas.

#### *Established ratings - ratings curve.*

For players with established ratings, I recommend calculating the expected win percentage for each game using:

$$1 / (1 + \text{Exp}(\text{AbsoluteValue}(\text{PlayerRating} - \text{OpponentRating}) / -313))$$

where Exp(x) is a widely available function which gives the value of e raised to power of x (i.e. the inverse of the natural logarithm). The value 313 is as explained above. The calculation of expectancies for players with established ratings would otherwise be as in the current system.

#### *Established ratings - k-factor.*

I recommend amending the formula for calculating the k-factor to:

$$(3000 - \text{currentRating}) / 1000 * \text{ratedGames}$$

which is equivalent to half of the current k-factor, multiplied by 1/10th of the number of games that the new established rating is being calculated for. I.e. playing 20 rated games in a tournament would give the same k-factor as the current system, playing 10 games would give half the current k-factor (up or down).

#### *Accelerator and feedback points.*

I recommend that points gained in excess of 5 points per game played be doubled, and opponents of anyone in receipt of accelerator points receive 0.05 points per accelerator points, per game played against the player who received them. Very roughly, if someone in a 14-game double round robin exceeds their expectancy by around 4 games, they would receive around 25 accelerator points, and each of their opponents

---

## RATINGS SYSTEM REVIEW - IMPLEMENTATION GUIDE

would get 2 or 3 feedback points (based on having played the player twice). I recommend that these rates of accelerator and feedback points be reviewed after the first two years of operation of the new system.

Accelerator points would not apply to the rating achieved by a new player in their first NZ tournament, as there is no previous rating to compare it to.

### *Participation points.*

I recommend that 'participation' points be awarded, partly to combat ratings deflation (as higher-rated players leave the system and new players join the system), and partly to recognise that a new rating is more reliably comparable to other ratings than the rating of someone who may not have played for a long time.

I recommend trialling a rate of 1 participation point for each 3 games played in a tournament (i.e. 6, 7, or 8 games played would get 2 points added after the new rating is calculated, 12, 13, or 14 games played would get 4 points, etc.).

### *New, provisional, and historic ratings.*

I recommend that any player who is provisionally-rated at the beginning of a tournament may be placed in a different grade to that indicated by their rating (if the player and tournament organiser agree on a suitable grade), in the same way that new players and players with historic ratings may already be re-graded.

I recommend that the current method of re-rating players with historic ratings if they are played in a different grade be discontinued, in light of the adoption of accelerator, feedback, and participation points.

## Implementation of Transitional Measures

### *Provisional ratings.*

For players listed as having provisional ratings in the end-of-year ratings (i.e. who have played fewer than 35 games in the current system), if they have played at least 30 games, their current provisional rating will become an established rating for the start of the new system. If they have played fewer than 30 games, their provisional ratings after each of their prior tournaments will be used in the new system as if it was a partial provisional rating under the new system.

For example, take a player who has played a 7-game tournament, and a 13-game tournament under the current system, and then plays a 7-game tournament under the new system. The overall calculation will be to add 7 times the provisional rating after the first tournament to 13 times the provisional rating after the second tournament, and then add the 7 per-game provisional ratings from the third tournament separately. Then divide the whole lot by 27. Subsequent tournaments can then revert to the new system - i.e. add 27 times the initial provisional rating to the new per-game ratings, then divide by the new total games.

## Other implementation details not covered above

- i) ~~That the above proposed changes and transitional measures be adopted for tournaments starting with the first tournament next year (2017).~~
- ii) ~~That I produce a cross-platform implementation of the new system, that can be used to easily administer it.~~
- iii) ~~That the new system be reviewed after two years of operation.~~

---

## RATINGS SYSTEM REVIEW - IMPLEMENTATION GUIDE

(Please note that some formulas above omit the rounding required to give whole number ratings. This is solely for the sake of clarity—any rounding that is part of the current system will carry over to the new system, as appropriate.)

I would particularly like to thank Chris Handley, Ed Okulicz, Glenda, Olivia, and those who attended the June management meeting for their input into various stages of this review.

