

Review of the 2019 changes to the NZASP Ratings System

Final Report

Prepared by: Steven Brown (Rating System Manager)

Assisted by members of the Ratings Review Panel:

Dylan Early (convenor), Steven Brown, Patrick Carter, John Foster, Alastair Richards

8 October 2021

REVIEW OF THE 2019 CHANGES TO THE NZASP RATINGS SYSTEM

Background

The executive summary of the recommendations of the 2018 rating system review, as accepted by the Management Meeting on Saturday 26 January 2019, included recommendation 6, that:

The recommendation is to implement the new system at the beginning of this year (2019) and review it after two years unless indicated earlier.

This review was postponed last year (2020), and due to pandemic suspensions, two full years' worth of games would not have been completed in 2019-2020. Two full years of games had been completed by the middle of this year (2021), so the review is being completed in time for any recommendations to take effect from 1 January 2022.

Analysis

This report will consider each recommendation of the report of the 2018 ratings system review, and discuss whether the recommendation has been followed, the effect this has had, and any other issues that have arisen relating to that recommendation.

The report will conclude with any recommendations arising from the review.

Ratings curve

Recommendation for 2019

Keep the 2017 system curve. Further, more rigorous, analysis of the 18 years of the 1999 system, and of the 2 years using the new curve, show the same results as the analysis undertaken in 2016 - that it is the best fit of any of the curves that have been used or considered for use by major rating systems. Some other curves, not used in or considered by other ratings system might be an even better fit, but this is outweighed by the benefits of continuing to use the same curve as WESPA (and virtually the same as NASPA).

Discussion

Examination of tournament results since the beginning of 2019 has shown that players continue to have a roughly equal chance of gaining or losing ratings points, wherever in a grade their rating places them. The new ratings curve introduced in 2017 still appears to be significantly fairer than the original curve it replaced.

Multiplier or k-factor

Recommendation for 2019

Change the k-factor formula to: **$((3000 - \text{rating}) / 100) * 1.5$**

This is equivalent to the k-factor for a 15-game tournament under the 2017 system, and thus provides the best continuity between the 2017 system and what we are recommending.

Discussion

There have been no complaints that number of points gained or lost per game's difference between wins and expected wins is either too large or too small, so this change appears to be working well.

Participation points

Recommendation for 2019

Remove entirely, but investigate in a future review the possibility of pairing their reintroduction with some other method of adjusting the overall level of ratings.

Discussion

Player's ratings have become noticeably more stable since the beginning of 2019, which would indicate that participation points were one factor in the ratings inflation in 2017-18, and that removing them has helped to stabilise the ratings system.

Treatment of New, Provisionally-rated, and Historically-rated players

Recommendation for 2019

We recommend changing the system for provisional ratings to a system where a player is still classed as provisional until they have played at least 30 NZ tournament games, but where the iterative performance rating mentioned above is calculated only for their very first tournament, whereafter they are treated by the rating system as having an established rating. This would both greatly simplify this aspect of the ratings system, and would mean that all grades could have exact pre-game expectancies easily calculated, except those that include brand-new players. It would also remove the need to implement the second post-World Seniors fix.

The new system would then run the risk, however, of a new player being placed in an inappropriately low grade, and not then having the opportunity of further provisionally-rated games to improve their initial rating. We recommend that this be addressed by keeping acceleration points (and feedback points for opponents) for such players, and for players with a historic rating (but with a higher threshold, as discussed below).

Discussion

Calculating a player's first rating by using the standard method for calculating a rating repeatedly, and replacing the players initial rating with the final rating from the previous application of the method each time, until the final rating is the same as the initial rating, has proven to work very well. The new player's first rating will, if they win all their games, be somewhat above the highest rating in the grade in which they played, and, if they lose all their games, somewhat below the lowest rating in the grade.

The risk of a player being placed in an inappropriate grade for their first tournament remains. A strong player who is new to the NZ rating system might be placed in too low a grade, and gain too low an initial rating even if they do well, or the playing strength of a new player might be overestimated, in which case placing them in too high a grade would lead to them gaining too high an initial rating, even if they did badly. The second case was not addressed by the retention of accelerator/feedback points for provisionally-rated players.

These cases are similar to the case of historically-rated players in that it is possible for them to start a tournament with a manifestly inappropriate rating, and hence grading. In the case of a player with a historical rating, this may be because the player's playing strength has gotten demonstrably better or worse since they last played in a NZ-rated tournament, or because the playing strength represented by a particular ratings value may have drifted over time. A 2013 amendment to clause 15.2 of the NZASP Constitution put forward by the Ratings Officer, which became rule 2.3 in the Rules of Play when they were separated out in 2016, provided that:

REVIEW OF NZASP RATINGS SYSTEM CHANGES - FINAL REPORT

“2.3. Where a player has not played a rated game in New Zealand for at least two years (i.e. they have a historic New Zealand rating, but no longer appear in the Ratings list), the player and Tournament Organiser may agree that the player should be placed in a higher or lower grade than indicated by their historic rating.

In these cases, the player's new rating should be calculated in such a way as to not unduly advantage or disadvantage the other players in the grade.”

Extending this to players who are classed as provisionally-rated under the 2019 system would address the risk of new players getting an inappropriate rating from their first tournament. Calculating the ratings for such players using the current method for new players would meet the requirements of the last sentence in Rule 2.3 above, and would allow Acceleration and Feedback points to be finally removed entirely from the system.

Acceleration and Feedback points

Recommendation for 2019

Remove Acceleration and Feedback points (except as recommended for simplified provisional ratings). For provisional and historic players, increase the threshold to 3.5 games, making them slightly harder to earn.

Discussion

The removal of Acceleration and Feedback points for most players has both greatly simplified the system, and removed the problem of such points being earned by higher-ranked players, who arguably didn't need them. Removing them entirely would finish the job of making the system simpler both to implement and administer, and to explain.

Other matters—World Seniors 2017

Recommendation for 2019

We believe that this tournament should not have been rated in the NZ system, but consider the question of which tournaments should be rated is outside the scope of both this review and the responsibilities of the Ratings Manager, and should be better defined in the Constitution.

Discussion

The question of which tournaments should be rated in the NZ system remains unaddressed, and remains outside the scope of both this review and the responsibilities of the Ratings Manager.

Other matters—Range of ratings in the system

Recommendation for 2019

There is sufficient disagreement about the causes, severity, and appropriateness of recent rises in the ratings of top-rated players that we consider that this matter should be left for further consideration by a subsequent review, when the system has had time to adjust to the new (2017) ratings curve and there is a further year or two of data to analyse.

Discussion

Given that individual ratings have on average been much more stable since the beginning of 2019 than they were during 2017-18, the causes of the rises in the ratings of top-rated players during 2017-18 seem to have

been some combination of the system's initial adjustment to the new ratings curve, the effect of accelerator/feedback points, and the emergence or improvement of several top-rated players during that time. How much each factor contributed to the phenomenon would be difficult to untangle.

Recommendations

Recommendation 1

That, effective from 1 January 2022, the Rules Committee should replace clause 2.3 of the rules of play with the following:

"2.3. Where a player either:

- i) has played at least one rated game in New Zealand, but fewer than 30 (i.e. they have a provisional New Zealand rating); or
- ii) has not played a rated game in New Zealand for at least two years (i.e. they have a historic New Zealand rating, but no longer appear in the Ratings list);

the player and Tournament Organiser may agree that the player should be placed in a higher or lower grade than indicated by their provisional or historic rating.

Any player thus re-graded will be rated as if they are a new player who has been placed in that grade."

Recommendation 2

That, from 1 January 2022, the ratings system should be operated in accordance with new clause 2.3 of the Rules of Play.

Recommendation 3

That the Rating System Manager and members of the Ratings Review Panel should continue to monitor the functioning of the ratings system, but that it should only next be reviewed when, in the opinion of the NZASP Executive, there is sufficient demand for this to occur.