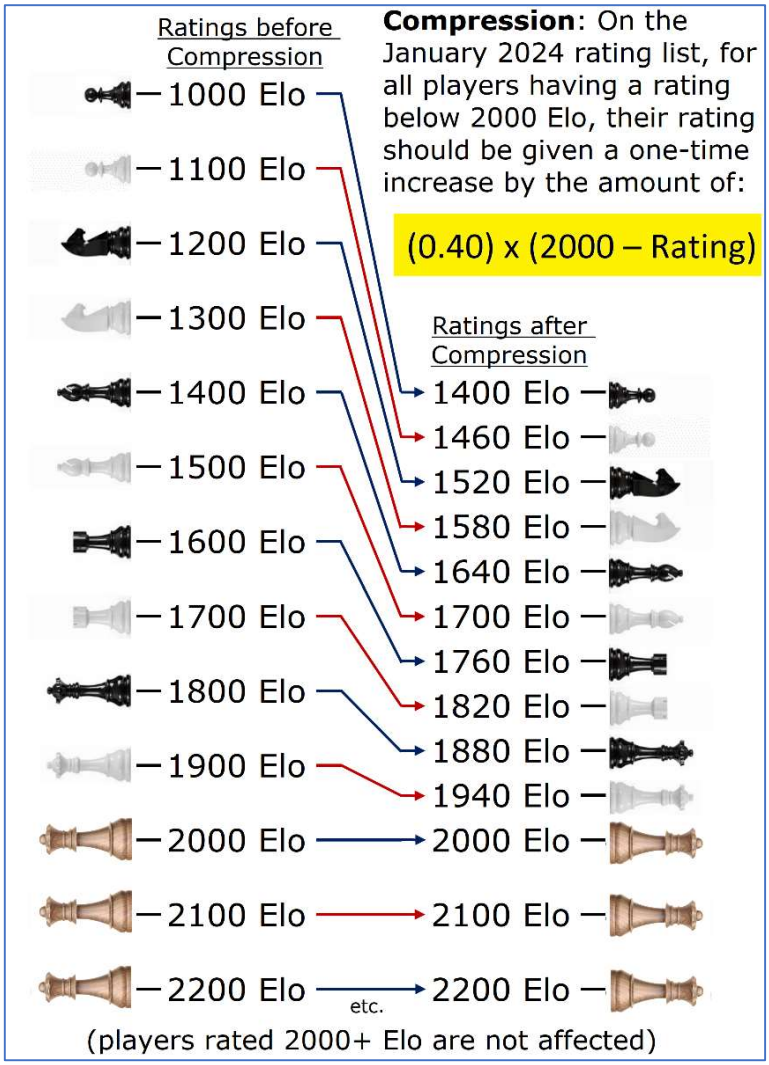
Sonas Proposal: Repairing the FIDE Standard Elo rating system

-- by Jeff Sonas (20 July 2023)

Introduction and Executive Summary

To put it bluntly, the past decade has brought extreme rating deflation to the FIDE Standard Elo rating system. Player ratings are spread apart far more than they should be, and it's getting worse every year. FIDE ought to take strong steps to reverse the effects of this deflation, and to prevent it from happening again. I propose two corrective measures to accomplish these joint goals.

The first corrective measure, which I am calling the **Compression**, will bring immediate and powerful improvement to the accuracy of the ratings. All players rated below 2000 Elo, which is the bottom 85% of rated players, should be given a one-time rating increase, on the January 2024 rating list, using a simple formula as highlighted in yellow in the accompanying graphic.



For example, look at the black pawn on the left, representing anyone who has a rating of 1000 Elo before the **Compression** happens. Such a player would get an increase of 400 Elo points and would end up with a rating of 1400 Elo, as you can see from the black pawn on the right.

As another example, look at the white bishop on the left, representing anyone who has a rating of 1500 Elo before the **Compression** happens. They would get an increase of 200 Elo points and would end up with a rating of 1700 Elo, as you can see from the white bishop on the right.

Also note that there are no discontinuities in the formula; for example, there is no advantage or disadvantage to being just over 1500 or just under 1500. If you had 3 players with ratings of 1495, 1500, and 1505, respectively, they would receive one-time rating increases of 202, 200, and 198, respectively. The chess pieces in this diagram are provided just to help you follow the lines and to help illustrate what would happen for players at various places on the rating list.

Players rated just below 2000 Elo would get almost no increase at all. And for all players having a rating of 2000 Elo or higher, their rating would not be directly affected by this. Note that this change would preserve the current order of everyone on the rating list; it would only adjust their spacings between each other. And finally, it would leave nobody having a rating below 1400 Elo.

The second corrective measure, the **Calculation Improvements**, tries to ensure that the rating deflation will not return, by introducing some mild inflationary effects into the rating regulations. This includes raising the minimum rating from 1000 Elo up to 1400 Elo, restoring the 400-point-rule back to its earlier state so that it can apply multiple times in an event, and adjusting the formula used for calculating initial ratings so that it incorporates two additional draws against hypothetical 1800-rated opponents.

I appreciate that these are very significant steps to take, and I wouldn't expect anyone to support them without understanding why they are necessary. That is why I am not just proposing these measures, but also justifying them, as part of this proposal. The corrective measures are described in more detail in Part III, below, but first it is necessary to give some background and to describe the problem in detail.

Part I – Background

FIDE maintains an official worldwide chess rating list, based on the Elo rating system, for all reported games played at standard time controls. Each month, player ratings are updated, based on all the chess tournament results reported to FIDE that month, and the updated rating list is published online by FIDE. The Elo system ensures that players who exceeded their expected score will see their ratings increase, and players who fell short of their expected score will see their ratings decrease. New players qualify for the list once they have played enough games against existing rated players, and so the rating list grows larger each month.

The first official FIDE rating lists in the early 1970's included only players having an Elo rating above 2200. This restriction stayed in place for twenty years, while the rating list grew twenty times bigger, from 600 players in the early 1970's up to 12,000 players in the early 1990's. With the lowering of the minimum rating to 2000 Elo in January of 1993, FIDE began a new twenty-year phase of allowing progressively lower-rated players onto the rating lists. The minimum was dropped again to 1800 Elo in October 2001 and then incrementally to 1600, 1400, and 1200, until the current minimum of 1000 Elo was established in August 2012.

This gradual reduction of the minimum rating greatly increased the number of FIDE-rated chessplayers worldwide. For example, there were fewer than 40,000 players on the January 2003 rating list, compared to more than 400,000 players on the January 2023 list. Here is a graphic showing how the distribution of rated players (grouped together into 100-Elo-point ranges) has changed across those twenty years:

However, the reductions of the minimum rating had a very problematic side effect. Many younger/weaker players receive very low initial FIDE ratings now at a very young age, during the stage of their career when their chess skill is improving most rapidly, and this combination of low initial ratings and rapidly-improving players has ultimately produced a strong deflationary pressure upon the overall rating system for the past decade.

Twenty or thirty years ago, we were all worried about "rating inflation", as we watched more and more players rated 2700+ filling up the top 100, but we are long past the era of rating inflation. In fact, that inflationary era was probably already in the rear-view-mirror by the time of that (purple) January 2013 list we just saw. For many years now, "rating deflation" has actually been the greater problem, although you won't notice it just from watching the monthly top 100 lists.

In 2014, FIDE enacted measures designed to counteract rating deflation in the standard rating list, by increasing most K-factors and also maintaining high K-factors for weaker junior players even after they had played many rated games. The higher K-factors would cause those players' ratings to increase more quickly when they outperformed their rating expectation. The hope was that the ratings of improving juniors would therefore more easily keep up with their improving skill, and so their upcoming opponents would not be losing so many rating points when facing them. However, those 2014 measures have not proven sufficient to counteract the deflationary pressures.

Part II – The Deflation Problem

Although the Elo system is still functioning reasonably well for grandmaster-level players, the fact remains that 99% of rated players have ratings below 2400 Elo, and there are some very serious problems with their ratings.

The active rating pool doubled in size during the seven years from 2013-2020, thanks mostly to a large influx of weaker junior players, many of whom were immediately and significantly underrated with their very first rating. As these players started playing with greater skill than their own ratings would predict, the natural functioning of the Elo system started transferring a very large quantity of rating points to these new players out of the existing rating pool.

It's relatively easy to understand what "rating inflation" looks like – players get new ratings that are too high (relative to the existing rating pool), and they start losing their excess rating points into the pool of established players, who themselves now become slightly overrated, and then those players (in turn) slightly underperform their ratings and spread out the excess points even further… and over time those excess ratings points spread throughout the entire rating system.

It's a little more challenging to visualize what "rating deflation" looks like. Rather than an "excess" of rating points, it is instead a "deficiency" of rating points, but everything else works similarly. So these new players enter the rating system strongly underrated (relative to the existing rating pool), with a large "deficiency" of rating points. As they overperform their low ratings, and pull rating points away from their opponents, those opponents in turn acquire a "deficiency" of rating points (they become a bit underrated), and so on throughout the entire rating pool. Ultimately everyone's rating gets pulled down, although the severity will vary depending on everyone's degree of connection to new players (e.g., X plays Y who plays Z who plays a new player). Grandmasters are the ones furthest removed from new players, and therefore most shielded from the effect.

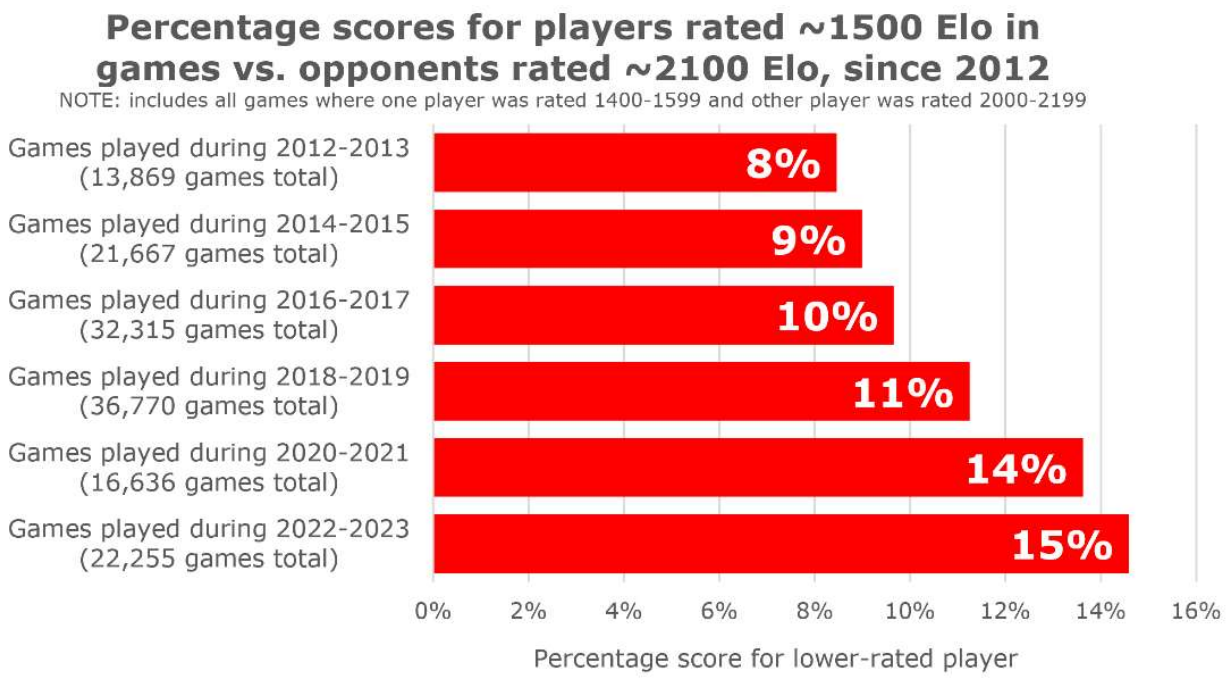
Everything I just described about "rating deflation" has really happened, during the past decade, to the FIDE Standard Elo rating system. New players have entered the system, hundreds of thousands of new players, many of them entering with a large deficiency of rating points, and their overall impact has been to pull down ratings of the established rating pool.

This deflationary effect wouldn't have been too problematic if the rating pool had been a closed group, because the Elo system would eventually find a proper equilibrium, though perhaps one with everyone having slightly lower ratings than before. However, the rating system never had time to catch its breath and establish such an equilibrium. There were always lots of additional new players entering the system each month, getting initial ratings that were far too low, and immediately pulling more rating points away from established players, even from junior opponents who themselves had quite recently been the new and underrated players! It is a constant deflationary pressure that is being applied even now, and it's getting worse every year, and it will not get better until significant counter-measures are implemented.

We might wonder if the COVID-19 shutdowns had actually helped this problem, because the influx of new players dropped down to essentially zero. However, the Elo system still didn't have time to catch its breath and find a proper equilibrium, because ratings can't change unless games are played, and there were essentially zero games being played among rated players during the shutdowns. So the Elo system ground to a halt during those times.

In fact, things got even worse for the FIDE Elo system with the onset of COVID-19 and the drastic reduction in over-the-board events. The lack of rated events did not keep all those weaker junior players from improving their chess skill, creeping ever closer toward the playing strength of master-level players. But the shutdown did severely restrict the ability of the Elo rating system to keep pace with the advancement of all those weaker junior players. Because of this, it seems quite plausible that a 1500 Elo rating today, for example, reflects a closer proximity to master-level skill than a 1500 Elo rating five or ten years ago would have represented.

Let's see if we can investigate that last point objectively, rather than engaging in theoretical speculation. For example, what happens when 1500-rated players face 2100-rated players in today's chess world? And is it any different from what happened five or ten years ago? Let's look at the strong empirical evidence from more than 140,000 rated games played since 2012, between players rated near 1500 Elo and opponents rated near 2100 Elo. In fact, the trend is quite clear:



We can say confidently that the gap in playing strength has greatly shrunk during the past decade between players in these two groups (~1500 Elo and ~2100 Elo) on any given rating list. This is absolutely not supposed to happen in the Elo system – a difference of X Elo points is always supposed to correspond (on average) to the same Y percentage score. Nevertheless, it is easily seen that such players are far closer in playing strength today, than they were five or ten years ago.

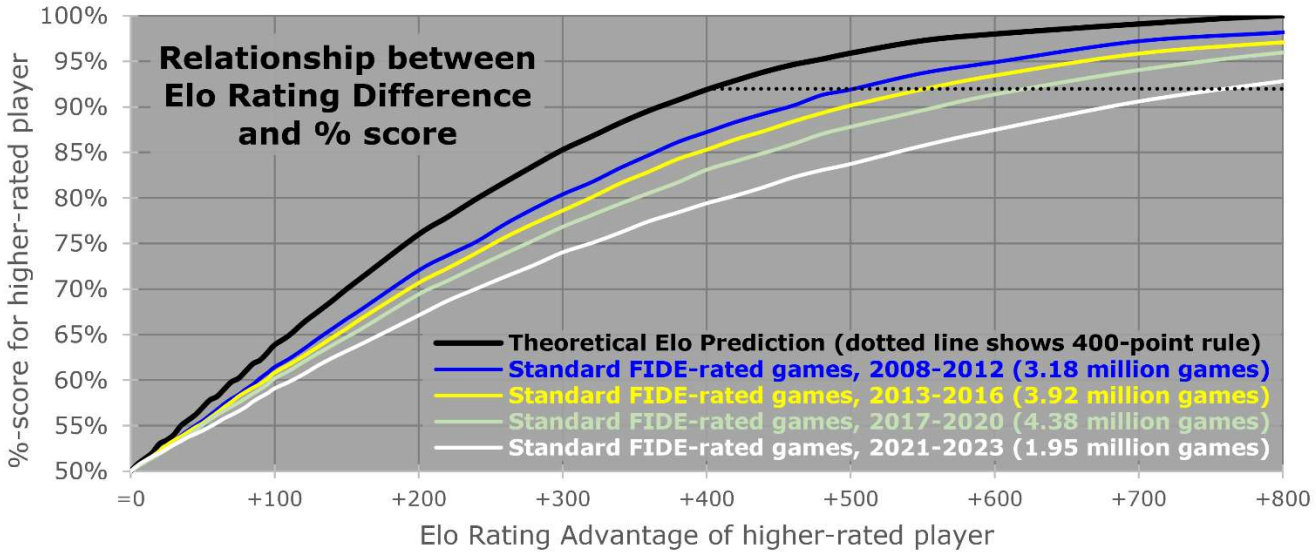
Ten years ago, the lower-rated players were scoring 8% in such games. More recently, they've been scoring 15% instead. Today's performances are what you would expect from players who are separated in strength by 300 Elo points, rather than the 600 Elo points suggested by the rating list. And by the way, the Elo table would tell us that the lower-rated players should score about 2% in these games.

The spacings between ratings simply do not have the same meaning that they once did. These two groups of players, showing strong evidence of being 300 Elo points apart in actual playing strength, are instead separated by 600 Elo points on the rating list. This is why the term "rating expansion" might actually be a more apt description than "rating deflation". In fact, the whole pool of ratings (below master-level) is stretching out each year, in the downward direction. And thus a rating gap of 600 Elo points on today's rating list does not correspond to the same difference in playing strength that it did just a few years ago, because ratings are now more stretched apart than they should be.

And of course, there is nothing special about a rating gap of 600 Elo points. This problem is present for all numerical rating gaps, small or large. Let's look at a far larger range, from a gap of 0 Elo rating points where each player has an expected score of 50%, all the way up to a gap of 800 Elo points where the higher-rated player essentially should score 100% according to the Elo tables. That range covers more than 99.5% of all rated games.

The following graphic shows the relationship between Elo rating difference and percentage score, across all FIDE-rated games (at standard time controls) since 2008 where the players were within 800 Elo points of each other. The horizontal axis indicates the Elo rating advantage for

the rating favorite, and the vertical axis indicates the percentage score for that player. The black curve shows the rating favorite's expected percentage score, based on looking up their rating difference in the Elo table. So you can see from the black curve that the higher-rated player has an expected score just above 50% if their rating advantage is small, up to an expected score of 85% with a +300 rating point advantage, and all the way up to essentially 100% with even larger rating advantages.



Other colors show the empirical results, namely the rating favorites' average scores in actual rated games, when they really did have those rating advantages. This shows all standard FIDE-rated games at these rating differences from games played in 2008-2012 (blue), games played in 2013-2016 (yellow), games played in 2017-2020 (green), and finally, games played in 2021-2023 (white). This one image summarizes more than 13 million rated standard games played over the past 15 years. And unfortunately, it paints quite a gloomy picture.

If the rating system had been functioning well across all these years, then the colored curves would perfectly overlay/intersect the black curve. However, all of the colored curves are shallower than the black curve, indicating that the rating favorites are consistently underperforming their ratings, whatever the size of their rating advantage might be.

Even ten or fifteen years ago, the situation wasn't great. The blue curve tells us that rating favorites were consistently underperforming in the games from 2008-2012. But then things got even worse, as we see with the yellow curve (2013-2016) being lower than blue, the green curve (2017-2020) being lower than yellow, and finally the white curve (2021-2023) being lowest of all. The rate of decline even seems to be accelerating.

We have never been further away from that black curve, that ideal relationship between Elo rating difference and percentage score, than we are with today's rating system and today's ratings. The white curve is the shallowest of them all, meaning that rating favorites are underperforming by more than they ever have before. And of course this is not just a theoretical oddity. Indeed, players lose rating points in direct proportion to how far their average percentage score falls below the black curve. So in practice, this trend means that rating favorites are losing more and more rating points per game played, each and every year.

In my opinion, the situation is worsening because we constantly see new players getting initial ratings that are way too low, and then when they play rated games against established players,

they are pulling those established ratings downward by taking rating points away from the established players. And this ultimately is stretching out the entire rating pool, leading to ratings that overstate the true differences in playing strength among players.

We should also acknowledge that this effect is not happening only among weaker players, or only among stronger players. Rather, it is present throughout the entire rating pool. We can double-check this by constructing a "summary crosstable" spanning several entire years of data. A traditional crosstable represents one event with individual players shown in each row and each column. But this "summary crosstable" would show the performance of entire rating groups (grouped together in intervals of 100 Elo points) against other entire rating groups, across many tournaments or even many years.

For example, let's make a summary crosstable for the games played between 2008 and 2012, with players assigned into groups depending on the first two digits of their Elo ratings. So we would have players rated 1200-1299 in one group, players rated 1300-1399 in another group, and so on. Between 2008 and 2012, players rated 1200-1299 Elo played a total of 726 games against opponents rated 1300-1399 Elo, with the lower-rated players expected to score 37% in those games and actually scoring 38%, so they exceeded their expected percentage score by +1%. And conversely, the 1300-1399 players scored 1% below expectation in the same games.

We can show this in a summary crosstable, where we color-code the cell backgrounds with a blue-red gradient, so overperformance is shown in blue as positive numbers, underperformance is shown in red as negative numbers, and bigger deviations from expected %-score are shown as darker blues and reds. A difference in 1% from expectation is quite small, so the red and blue backgrounds are very pale for that first example. Differences of 2% or 3% from expectation are progressively darker. Ideally every cell would be "0%" and a white background, if everyone were performing in line with their expectation.

opponent / player	vs. 1200-99	vs. 1300-99	vs. 1400-99
1200-99	X	**+1%**	**+3%**
1300-99	**-1%**	X	**+2%**
1400-99	**-3%**	**-2%**	X

We can find the 1200-99 group in the leftmost column, and read across to see their slight +1% overperformance (pale blue) against opponents in the 1300-99 group and a bigger +3% overperformance (medium-light blue) against opponents in the 1400-99 group. In both cases, the players rated 1200-99 were overperforming against their higher-rated opponents.

Similarly, we can find the 1400-99 group in the leftmost column, and read across to see they had a -3% underperformance (medium-light red) against opponents in the 1200-99 group and a -2% underperformance (light red) against opponents in the 1300-99 group.

The diagonal is filled with X's, since we aren't showing what happens when a group plays against itself (the numbers would always be 50% expected and 50% actual). And keep in mind that the "lower-left" region (the cells below and to the left of the diagonal) represent the performance of rating favorites during games from 2008-2012, and the "upper-right" region (the cells above and

to the right of the diagonal) represent the performance of rating underdogs during those same games. Based on our earlier understanding that even in 2008-2012, rating favorites were falling short of their predicted scores, we can expect to see mostly red cells in the lower-left and mostly blue cells in the upper-right, and at least this small example does follow that trend.

Of course, the picture gets far more interesting and informative if we show, not just the groups rated below 1500 and how they all did against each other, but actually all the different groups against all other groups. We will lump together all the players with 2500+ Elo into their own single group so we can show significant amounts of data for them, rather than trying to show separate data for 2500-99, 2600-99, etc. Here is the full picture for games played during the five-year span from January 2008 through December 2012:

2008-2012	opponent	vs.	vs.	vs.	vs.	vs.	vs.	vs.	vs.	vs.	vs.	vs.	vs.	vs.	vs.
	player	1200-99	1300-99	1400-99	1500-99	1600-99	1700-99	1800-99	1900-99	2000-99	2100-99	2200-99	2300-99	2400-99	2500+
+4.5% in 6,865 games	**1200-99**		+1%	+3%	+6%	+6%	+7%	+4%	+4%	+2%	+2%	=0%	=0%	=0%	=0%
+4.8% in 25,544 games	**1300-99**	-1%		+2%	+4%	+7%	+7%	+4%	+5%	+4%	+4%	+1%	=0%	+2%	=0%
+5.0% in 60,420 games	**1400-99**	-3%	-2%		+3%	+5%	+7%	+8%	+7%	+4%	+3%	+1%	+1%	+1%	=0%
+4.9% in 135,409 games	**1500-99**	-6%	-4%	-3%		+3%	+6%	+7%	+8%	+6%	+5%	+3%	+2%	+1%	=0%
+4.0% in 241,595 games	**1600-99**	-6%	-7%	-5%	-3%		+4%	+6%	+8%	+7%	+6%	+4%	+2%	+2%	+0%
+2.7% in 385,143 games	**1700-99**	-7%	-7%	-7%	-6%	-4%		+3%	+6%	+7%	+6%	+5%	+3%	+1%	+1%
+1.4% in 540,108 games	**1800-99**	-4%	-4%	-8%	-7%	-6%	-3%		+3%	+5%	+6%	+4%	+4%	+2%	+1%
+0.0% in 699,217 games	**1900-99**	-4%	-5%	-7%	-8%	-8%	-6%	-3%		+3%	+4%	+4%	+3%	+2%	+0%
-1.0% in 801,405 games	**2000-99**	-2%	-4%	-4%	-6%	-7%	-7%	-5%	-3%		+2%	+3%	+3%	+2%	+1%
-1.5% in 779,484 games	**2100-99**	-2%	-4%	-3%	-5%	-6%	-6%	-6%	-4%	-2%		+2%	+2%	+2%	+1%
-1.3% in 644,501 games	**2200-99**	=0%	-1%	-1%	-3%	-4%	-5%	-4%	-4%	-3%	-2%		+2%	+2%	+1%
-1.3% in 453,537 games	**2300-99**	=0%	=0%	-1%	-2%	-2%	-3%	-4%	-3%	-3%	-2%	-2%		+2%	+1%
-1.1% in 315,497 games	**2400-99**	=0%	-2%	-1%	-1%	-2%	-1%	-2%	-2%	-2%	-2%	-2%	-2%		+1%
-0.8% in 213,368 games	**2500+**	=0%	=0%	=0%	=0%	-0%	-1%	-1%	-0%	-1%	-1%	-1%	-1%	-1%	

Remarkably, even 12-15 years ago there was no blue at all in the entire lower-left region. The ratings were stretched out too far even in 2008-2012, since the rating favorites were underperforming within all of these matchups having any reasonable amount of data, and even in the ones without much data, they are only holding their own at "=0%" in the white cells.

Okay, so that was 2008-2012. A very long time ago. Let's jump forward by more than a decade and generate this same type of summary crosstable again, but with more recent data, namely the games from 2021-2023. So first of all we need to include two additional rating groups (players rated 1000-99 and players rated 1100-99), since Elo ratings reach down lower these days than they did before 2012. So there's two more rows at the top, and two more columns on the left. We will focus on checking whether it's still all red in the lower-left, and all blue in the upper-right. What do we see?

2021-2023	opponent	vs.	vs.	vs.	vs.	vs.	vs.	vs.	vs.	vs.	vs.	vs.	vs.	vs.	vs.	vs.	vs.
	player	1000-99	1100-99	1200-99	1300-99	1400-99	1500-99	1600-99	1700-99	1800-99	1900-99	2000-99	2100-99	2200-99	2300-99	2400-99	2500+
+7.4% in 47,760 games	**1000-99**		+2%	+7%	+9%	+10%	+10%	+10%	+9%	+8%	+5%	+5%	+4%	+2%	=0%	=0%	=0%
+7.8% in 150,873 games	**1100-99**	-2%		+4%	+7%	+9%	+12%	+12%	+10%	+9%	+7%	+6%	+5%	+4%	+1%	+1%	=0%
+6.8% in 205,982 games	**1200-99**	-7%	-4%		+5%	+9%	+12%	+14%	+13%	+12%	+9%	+8%	+6%	+4%	+2%	+1%	=0%
+5.3% in 243,185 games	**1300-99**	-9%	-7%	-5%		+5%	+11%	+14%	+15%	+14%	+12%	+10%	+9%	+7%	+4%	+3%	+3%
+3.7% in 270,932 games	**1400-99**	-10%	-9%	-9%	-5%		+6%	+11%	+14%	+15%	+14%	+12%	+9%	+7%	+5%	+2%	+3%
+1.6% in 291,021 games	**1500-99**	-10%	-12%	-12%	-11%	-6%		+6%	+11%	+14%	+14%	+13%	+10%	+9%	+6%	+4%	+0%
+0.1% in 303,892 games	**1600-99**	-10%	-12%	-14%	-14%	-11%	-6%		+6%	+10%	+13%	+14%	+13%	+10%	+7%	+5%	+3%
-1.3% in 318,533 games	**1700-99**	-9%	-10%	-13%	-15%	-14%	-11%	-6%		+5%	+10%	+12%	+11%	+11%	+9%	+6%	+3%
-2.3% in 322,168 games	**1800-99**	-8%	-9%	-12%	-14%	-15%	-14%	-10%	-5%		+5%	+9%	+10%	+10%	+10%	+9%	+3%
-3.2% in 309,028 games	**1900-99**	-5%	-7%	-9%	-12%	-14%	-14%	-13%	-10%	-5%		+5%	+8%	+9%	+9%	+7%	+4%
-3.8% in 275,731 games	**2000-99**	-5%	-6%	-8%	-10%	-12%	-13%	-14%	-12%	-9%	-5%		+4%	+7%	+7%	+7%	+4%
-3.8% in 229,272 games	**2100-99**	-4%	-5%	-6%	-9%	-9%	-10%	-13%	-11%	-10%	-8%	-4%		+4%	+6%	+6%	+5%
-3.9% in 174,867 games	**2200-99**	-2%	-4%	-4%	-7%	-7%	-9%	-10%	-11%	-10%	-9%	-7%	-4%		+4%	+4%	+3%
-3.7% in 128,497 games	**2300-99**	=0%	-1%	-2%	-4%	-5%	-6%	-7%	-9%	-10%	-9%	-7%	-6%	-4%		+3%	+2%
-3.2% in 97,087 games	**2400-99**	=0%	-1%	-1%	-3%	-2%	-4%	-5%	-6%	-9%	-7%	-7%	-6%	-4%	-3%		+2%
-2.5% in 62,886 games	**2500+**	=0%	=0%	=0%	-3%	-3%	-0%	-3%	-3%	-3%	-4%	-4%	-5%	-3%	-2%	-2%	

Hmm, well not only is the effect still present everywhere, but in fact the red/blue shading is far darker, since many of the numbers are up to double-digits, indicating that in many cases the rating favorites are underperforming their expectation by 10% or even 15% per game! Things have gotten markedly worse. And in fact, we could have predicted this, if we'd remembered how much shallower that white curve was in the first graph (for the games from 2021-2023), even compared to how the situation had been only a few years previously.

The deflation has crept upward through the master level, and has even started to reach the grandmaster level. The elite players mostly face each other, and so this deflationary effect, that for years has propagated upward through the rating pool from new ratings being too low, has not fully hit elite players yet. But it's coming. Even now you can see red underperformance from the 2400-99 and 2500+ groups against opponents rated near 2000 Elo, where you did not see this 10-15 years ago. Even near the top, the differences in ratings do not mean what they once did.

However, let's set aside the grandmasters and only focus on the remaining 99% of FIDE-rated players, the ones having Elo ratings in the range from 2400 Elo down to 1000 Elo. The ratings of these players span 1400 Elo points from top to bottom. But if we carefully analyze the evidence of millions of recent chess games, checking the performance of various rating groups against other rating groups through data visualizations like these summary crosstables, and other types of analysis, we come to an inescapable conclusion.

The actual differences in playing strength across that group, those 99% of all FIDE-rated players, only span 1,000 Elo points! Not 1,400 Elo points. Which raises some very serious points. The lowest-rated players, the ones just over 1000 Elo? Perhaps they really should have ratings more like 1400 Elo! And the players rated around 1500 Elo? Perhaps they really should have ratings more like 1700 Elo! That's what the performance numbers would suggest.

The 2400-rated players are not performing a full 100 points better than the 2300-rated players, who in turn are not performing a full 100 points better than the 2200-rated players, and so on all the way down to the 1000-rated players. None of the rating favorites are able to reach the percentage score that their published Elo advantages would suggest. That's why the lower-left region is so strongly red in that last picture.

The FIDE standard Elo rating pool is stretched out too much - far too much - and there is very little incentive for any player to voluntarily face lower-rated opponents, since the likeliest outcome for the higher-rated player would always be a significant loss of rating points. And it's getting worse with each passing year. Honestly, I see no feasible course of action other than manually compressing the lower part of the rating list so ratings are not so spread apart, along with other measures to keep all this from happening again.

Part III – The Solution: Compression and Calculation Improvements

In recent months, I have been looking into this problem for FIDE. I developed a database tool that could replicate the past 15 years of monthly FIDE rating calculations, from 2008 to the present day, based upon the many millions of game results that were actually used for FIDE's official rating lists. This tool supports the various rating regulation changes that occurred over those 15 years, such as the progressively lower minimum ratings in 2009 and 2012, the changes to the K-factors in 2011 and 2014, the changes to the initial rating formula in 2014, etc.

During development of this tool, I verified that I could replicate the calculation of the actual historical rating lists quite precisely. But more importantly, the tool allows me to experiment with how the rating system would have behaved across all those years if different regulations had been in effect at the time. Obviously, any real changes we are contemplating would actually be deployed in the future, not in the past, but we can at least simulate what would have happened to ratings, based on real game data, if we had addressed this situation several years ago. This allows us to observe some of the likely long-term effects of any changes we are contemplating, as we calculate simulated ratings each month based on the games that actually did happen, and ultimately see how the ratings would look different today.

I have simulated a lot of different scenarios, in parallel with various discussions with FIDE Qualification Commission members, and I have settled upon my recommended solution to address the problems I just described. I have simulated scenarios where this recommended solution was deployed very recently, when the rating pool was quite similar to what we have today in mid-2023. This allowed me to get a clear sense of the expected short-term effects after the changes. Additionally, I have tried out a scenario where this solution was deployed in January 2017, allowing me to observe both the short-term effects and the long-term effects across the next 6+ years of simulated rating calculations. This gives us the best chance at estimating what would really happen to the FIDE rating pool in future years, if these changes were to go into effect soon. Admittedly, the onset of COVID-19 in 2020 does cause simulation challenges.

In any event, I propose that FIDE should undertake significant repairs to its Standard Elo Rating System by implementing two major changes, which I am calling the **Compression** and the **Calculation Improvements**. They are described as occurring in January/February 2024, because the **Compression** needs to happen 1 month before the **Calculation Improvements**, but they could just as well happen in October/November of 2023, or March/April of 2024. In any event, these repairs ought to be implemented as soon as possible. Here are my recommendations, which I believe to be FIDE's wisest course of action at this time:

III (a) Compression:

On the January 2024 rating list, for all players having a rating below 2000 Elo, their rating should be given a one-time increase by the amount of:

$$(0.40) \times (2000 - Rating)$$

The increase should be applied after the other calculations are already performed on the rating list for that month. It should happen for all active and inactive players rated below 2000 Elo, and it should happen whether or not they played any games for that rating period.

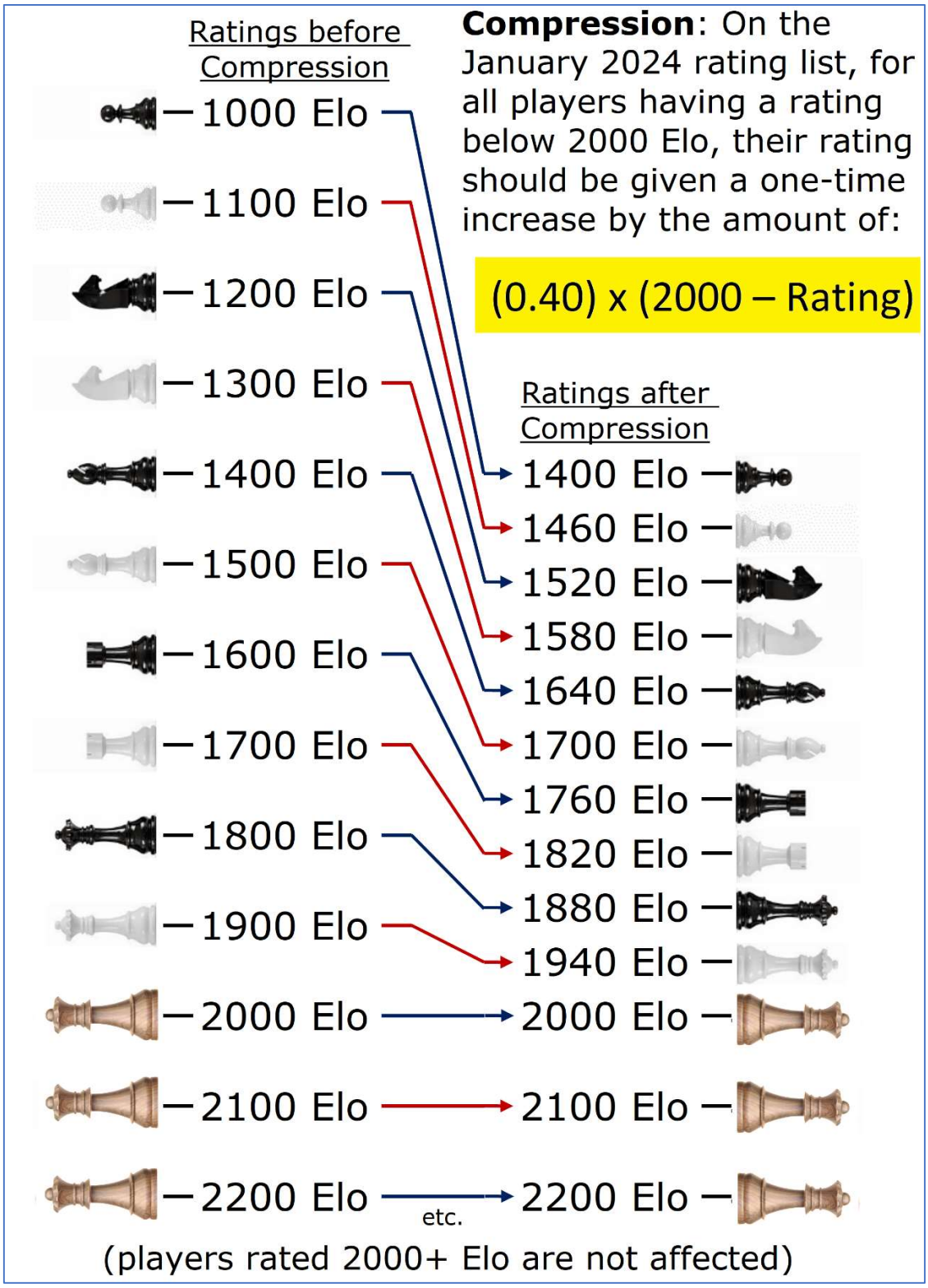
The **Compression** would take the group of players rated between 1000 Elo and 2000 Elo, which is the bottom 85% of rated players, and it would compress their ratings closer together so that they only span 1400 Elo to 2000 Elo instead. It would increase the ratings of all sub-2000 players by somewhere between 0 and 400 Elo points. Players rated just above 1000 Elo would get an increase of almost 400 Elo points, players rated near 1500 Elo would get an increase of about 200 Elo points, and players rated just below 2000 Elo would get almost no increase at all. And for all players having a rating of 2000 Elo or higher, their rating would not be directly affected by this. Note that this change would preserve the current order of everyone on the rating list; it would only adjust their spacings between each other. And finally, it would leave nobody having a current rating below 1400 Elo.

Here are some representative examples to further illustrate how the **Compression** would work:

Example 1A: After normal rating updates are calculated for January 2024, Player A has an Elo rating of 1200. Because this rating is below 2000 Elo, Player A receives a one-time rating increase of *(0.40) x (2000 – 1200) = 320 Elo points*. Therefore, Player A will have a rating of 1520 Elo on the January 2024 rating list.

Example 1B: Player B already has a rating of 1003 Elo and does not have any rated games for January 2024. Because this rating is below 2000 Elo, Player B receives a one-time rating increase of *(0.40) x (2000 – 1003) = 398.8 Elo points, rounded to the nearest integer of 399 Elo points*. Therefore, Player B will have a rating of 1402 Elo on the January 2024 rating list.

Example 1C: Player C is an unrated player who qualifies for an initial rating of 1900 Elo for January 2024. Because this rating is below 2000 Elo, and because the increase is applied after the regular calculations for the rating list, Player C receives a one-time rating increase of *(0.40) x (2000-1900) = 40 Elo points.* Therefore, Player C will have an initial rating of 1940 Elo on the January 2024 rating list.



Example 1D: Player D is an unrated player who does not yet qualify for an initial rating for the January 2024 rating list. Therefore, Player D continues to be unrated and will not ever receive a one-time rating increase, although they may eventually receive an initial rating through normal means, at some point in the future.

Example 1E: After normal rating updates are calculated for January 2024, Player E has an Elo rating of 2350. Because this rating is not below 2000 Elo, Player E's rating is not affected, and Player E will not ever receive a one-time rating increase. Therefore, Player E will have a rating of 2350 Elo on the January 2024 rating list.

III (b) Calculation Improvements:

Starting with the February 2024 rating list calculation, and for all future rating lists, the rating calculation algorithm should be changed in a few important ways:

(1) The minimum Elo rating will be 1400 Elo rather than 1000 Elo.
(2) The 400-point-rule will be restored to its previous state, with no restrictions on how many times it can be applied to a player during a single tournament.
(3) When adding up the results of games by an unrated player against rated opponents, for purposes of calculating an initial rating, the unrated player will be credited with two additional draws against hypothetical opponents rated 1800 Elo, in addition to their actual game results. All existing requirements would still apply, for instance requiring that players have a total of five or more actual games against rated opponents (not counting the hypothetical ones).
(4) When unrated players have a plus score in their games against rated opponents, their initial rating will be calculated the same way as for minus scores, being a performance rating based on their percentage score, rather than using the plus score multiplied by (K/2). However, the maximum initial rating that can be achieved from this formula will be 2200 Elo. If a player would get an initial rating higher than 2200 Elo, they will instead receive an initial rating of 2200 Elo.
(5) For unrated players who don't have a rating yet, all partial unrated performances in the past, achieved against rated opponents prior to the **Compression**, will be ignored and discarded. So all unrated players will get a "fresh start" at this point for getting a new rating.

Here are some representative examples to further illustrate how the **Calculation Improvements** would work:

Example 2A: Player A is an unrated player who scores 1 out of 4 in their first event against rated opponents, which is reported for the February 2024 rating list. Player A does not yet qualify for an initial rating, because they have not yet accumulated five results against rated opponents. The two hypothetical draws described in (3) above do not push them above the five-game-minimum. Therefore Player A will remain unrated on the February 2024 rating list.

Example 2B: Player B is an unrated player who scores 1 out of 6 in their first event, against rated opponents having an average rating of 1600 Elo, and this event is reported for the February 2024 rating list. Under previous regulations, a 17% score against 1600-level opposition would have yielded an initial rating of 1327 Elo. Instead, two hypothetical draws against 1800-level opposition are added to Player B's total results, for the purposes of this calculation. So instead, Player B is treated as though they scored 2 out of 8 (25%) against 1650-level opposition. Therefore, Player B receives an initial rating of 1457 Elo on the February 2024 rating list.

Example 2C: Player C is an unrated player who scored 1 out of 3 in their first event against rated opponents having an average rating of 1500 Elo, and this event was reported for the July 2023 rating list (before the **Compression**). Then in January 2024 (after the **Compression**) they score 0 out of 3 in their second event, against rated opponents having an average rating of 1700, and this event is reported for the February 2024 rating list. The July 2023 results are discarded because they happened before the **Compression**, and so the January 2024 results are viewed as their first result as unrated players. Since they had a zero score in that event, the results are discarded. Player C will remain unrated on the February 2024 rating list and will not have any results yet that count toward an initial rating.

Example 2D: Player D is an unrated player with no previous results and who has two events reported for the February 2024 rating list (after the **Compression**), scoring 1 out of 3 in their first event (against rated opponents having an average rating of 1500 Elo), and scoring 0 out of 3 in their second event (against rated opponents having an average rating of 1700). Just like with Player B, two hypothetical draws against 1800-level opposition are added to Player D's total results, for the purposes of this calculation. So Player D is treated as though they scored 2 out of 8 (25%) against 1650-level opposition. Player D receives an initial rating of 1457 Elo on the February 2024 rating list.

Example 2E: Player E has a 1430 rating on the January 2024 rating list and then scores 0.5 points in 7 rated games reported for the February 2024 rating list, where they were expected to score 2.5 points. From underperforming by 2.0 points with K=40, Player E should lose 80 Elo points. Because this would move them below the minimum rating of 1400, Player E becomes unrated for the February 2024 list (or they are reported with a rating of zero and then become unrated).

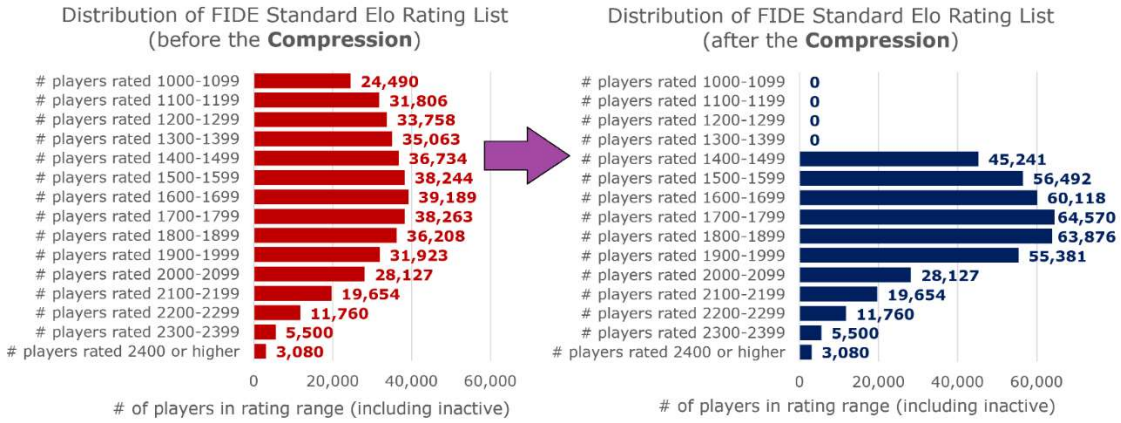
Example 2F: Player F is an unrated player who scores 5 out of 6 in their first event, against rated opponents having an average rating of 1600 Elo, and this event is reported for the February 2024 rating list. Under previous regulations, a +4 score against 1600-level opposition would have yielded an initial rating of 1680 Elo. Instead, two hypothetical draws against 1800-level opposition are added to Player F's total results, for the purposes of this calculation, and a performance rating is calculated. So instead, Player F is treated as though they scored 6 out of 8 (75%) against 1650-level opposition. Therefore, Player F receives an initial rating of 1843 Elo on the February 2024 rating list.

Part IV – Justifications and Expected Outcomes

Let us now try to project and predict the anticipated result if these fixes are implemented soon, as well as providing further rationale for them.

There would obviously be a very large change right away to player ratings, since approximately 85% of all rated players would get a rating increase of some sort from the **Compression**, and more than half of those increases would be 150+ Elo points.

As an example, if the **Compression** had been applied to the FIDE rating list in April 2023, here is a before/after comparison, showing how many players (including both active and inactive players) would have been in each 100-Elo-point range, with red showing the numbers before the compression, and dark blue showing the numbers after the compression:



Some key features of the **Compression** can easily be seen from this graphic. First of all, there would suddenly be zero players rated below 1400 Elo. All of the players who were previously rated 1000-1400 Elo would receive sufficient rating gains to push them above 1400 Elo. Further, one of the **Calculation Improvements** would be to immediately reset the minimum rating to 1400 Elo, and so there would continue to be no players rated below 1400 Elo, moving forward.

Although many players would receive rating gains, these gains would be smaller as you look higher in the rating pool, and in fact nobody's rating would be pushed above 2000 Elo by the compression. There would be the same 346,000 players rated below 2000 Elo both before and after; it's just that they would now span a range of 600 Elo points rather than 1,000 Elo points. And finally, none of the players rated 2000+ Elo would be directly affected by the compression. There would be the same 68,000 players as before, having the same ratings as before.

You might think, at first glance, that raising the minimum rating to 1400 Elo would exclude a large number of players in the future from getting ratings. After all, there are 125,000 players currently having ratings below 1400 Elo. What would happen to players who haven't yet achieved a FIDE rating, but are approximately as good as players rated 1250 Elo on today's rating list? Perhaps we should be more lenient and have a minimum rating of 1000 Elo or maybe 1200 Elo?

It's important to remember that all of the existing low-rated players would be getting significant increases to their ratings. Players who are currently rated 1250 Elo would end up around 1550 Elo as a result of the **Compression**. And so if there is an unrated player of comparable strength, facing opponents rated around 1550 Elo, then the unrated player's performance rating would probably be about 1550 Elo as well. They ought to have little trouble in getting an initial rating. To be honest, it might even become slightly easier for such players to get an initial rating, since another one of the **Calculation Improvements** would augment the results for unrated players by two draws against hypothetical 1800-rated opponents.

In fact, I think it would be unwise to do anything other than moving the minimum rating up to 1400 Elo. As a further illustration, let's consider players who are currently rated 1000-1050 Elo. These are today's lowest-rated players. After the **Compression**, these players would inhabit the range from 1400-1430 Elo. If we were to then allow new players to enter the rating list with ratings just above 1200 Elo, then those new players would be approximately 200 Elo points weaker than today's lowest-rated players. So going with a minimum rating of 1200 Elo would in fact be letting in players considerably weaker than those at the bottom of today's rating list. In contrast, using a minimum rating of 1400 Elo (after the ratings are compressed) would be more in line with today's standards. The only players who would be kept from getting a rating by the 1400 minimum would be players who likely wouldn't achieve a 1000 rating in the current system either.

The danger in going with a lower minimum rating than 1400 Elo, such as 1200 Elo, is that weak/improving players would get initial ratings that are likely far too low even for their current skill, and they would start pulling large amounts of rating points away from the established pool while their ratings catch up, which would be another strong deflationary effect. That is basically what has happened over the past decade with the low minimum rating of 1000 Elo. It would be a great shame to take all these steps to reverse the deflation of the last decade, only to encourage its return by making a choice that encourages more deflation.

You might also think, at first glance, that increasing the ratings of the weakest players by 400 Elo points is too large a leap, and that perhaps we should be satisfied with something like a "20% compression" rather than a "40% compression". This would result in only 200 Elo points being added to the lowest-rated players, and proportionally smaller increases for higher-rated players up to 2000 Elo. This is a reasonable reaction, but again I would caution against doing anything short of a 40% compression.

The proposed compression is based on detailed and extensive analysis that suggests the true performance level of the various rating groups. To take the players rated 1000-1100 Elo as an example, there are currently about 25,000 of them, making up 6% of the rating pool, and the evidence of their recent performance indicates they are really only 900-1000 points weaker than your typical 2400-rated player, despite being rated 1300-1400 Elo points lower. This conclusion is not just taken from direct results when players rated 1000-1100 Elo face 2400-rated opponents, since such matchups are quite rare and aren't very informative when they do happen. Rather, the conclusion is taken from simultaneous analysis of all the groups'

performances against all other groups, especially those rated nearby whom they face often. The compression would place the players currently rated 1000-1100 Elo into a much more appropriate Elo range of 1400-1460, in line with their evident strength.

The danger in going with a smaller compression than 40%, such as 20%, is that large blocks of improving players (particularly the weakest ones) would not be raised nearly high enough, and so they would continue to be strongly underrated and to pull large amounts of rating points away from the established rating pool when facing them. This would exert a significant deflationary pull on the overall rating pool, which is the very effect we're trying to counter with all of these proposed changes.

You might also be wondering why there is a need to add two additional draws against 1800-level opposition in the formula for new ratings. Previous analysis has indicated that in the current FIDE Elo system, players receiving relatively low initial ratings will tend to immediately start outperforming their rating (and gaining rating points), whereas players receiving relatively high initial ratings will immediately start underperforming their rating (and losing rating points). This is a strong indicator that there might be too much variability, too much spread, in the initial ratings that players have been receiving.

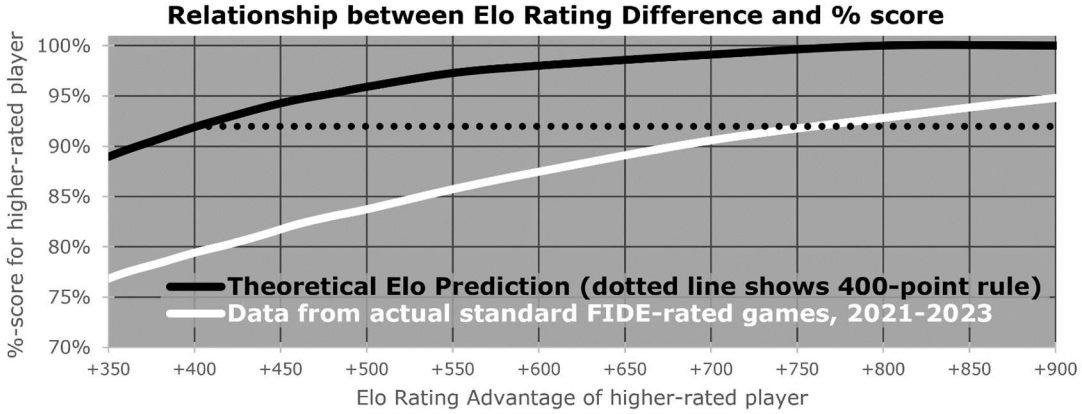
Furthermore, all indications point to the initial rating formula as being a primary contributor to the extreme rating deflation that we are currently experiencing. With much of the Elo system being a zero-sum game where rating points are merely exchanged between players, there is simply no other significant means by which large amounts of excess rating points, or conversely large deficiencies of rating points, can be introduced into the rating system. The initial rating formula is the key, being a heavy contributor both to the current problem and also to the proposed solution.

The Calculation Improvements, including two hypothetical drawn results against fairly strong opposition (1800 Elo), would build a mild inflationary factor into the initial rating formula to help counteract the natural deflationary factor that is inevitable when FIDE is allowing many weak/improving junior players to enter the rating pool. If some sort of counterbalance to the deflationary effect is not added like this, then we will likely be in the same boat again in a few years, even if we do implement a big one-time compression of the Elo rating pool.

By adding two hypothetical draws to each player’s performance while unrated, we will be pushing their percentage score in those games closer toward 50%, and therefore their initial rating will deviate less from their average opponent strength, and this will ultimately bring less variability to the initial ratings. And drawing against an 1800-rated opponent is likely a slight overperformance for most unrated players, and so these additional draws will "bake in" a small amount of planned improvement in the initial ratings of new players, so that their expected and natural improvement will not have quite so much impact upon the established rating pool as is the case today.

Special note about the 400-point-rule: *The 400-point-rule puts a cap of 92% on the expected score for a player having an extreme rating advantage (above 400 Elo points) in a rated game. FIDE recently implemented a regulation change so that a player cannot benefit from the 400-point-rule more than once per event.*

The relationship between Elo Rating Difference and Percentage Score has been shown already in this proposal, but let's zoom into one particular part of it, namely the part where the Elo advantage ranges from +350 up to +900. We will compare the black curve (showing the theoretical Elo prediction) against the white curve (the actual evidence of recent games from 2021-2023 where the Elo difference was between 350 and 900).



From looking at actual game data, we can see that currently, players with 400-point rating advantages are not even scoring 80% in such games, let alone 92%, and it takes something like a 650-700 Elo point rating advantage until we can reasonably expect the rating favorite to score as well as 90%.

Furthermore, in the current rating system, in those events where they are 400+ Elo point favorites in more than one game during the event, such higher-rated players would expect to be losing rating points in more than 70% of those events, even if the 400-point-rule were fully applied to all applicable games (not just one).

*Although it seems quite likely that the proposed **Compression** and **Calculation Improvements** will bring about drastic improvements, there is obviously no guarantee that this will happen. For the time being, in the current rating system, the deck is very heavily stacked against rating favorites. The above graphic tells us that even with the 400-point-rule applying fully, the rating favorites would still expect to lose rating points in games where the rating difference is 700 Elo points or closer.*

Therefore it seems counter-productive in the current environment to penalize extreme rating favorites for facing lower-rated players. My assessment is that the 400-point-rule is currently a useful tool for encouraging stronger players to face lower-rated players, and we should reinstate the behavior where the 400-point-rule applies to all games, even if a rating favorite already benefits from it elsewhere in the tournament. This also introduces a minor inflationary effect that slightly helps to counter-balance the deflation in the system.

Nevertheless, if these proposed changes do indeed have their desired effect, then I would recommend that either the 400-point-rule be weakened down to a 500-point or 600-point rule, or that it be removed altogether. Perhaps it can be re-assessed after a year.

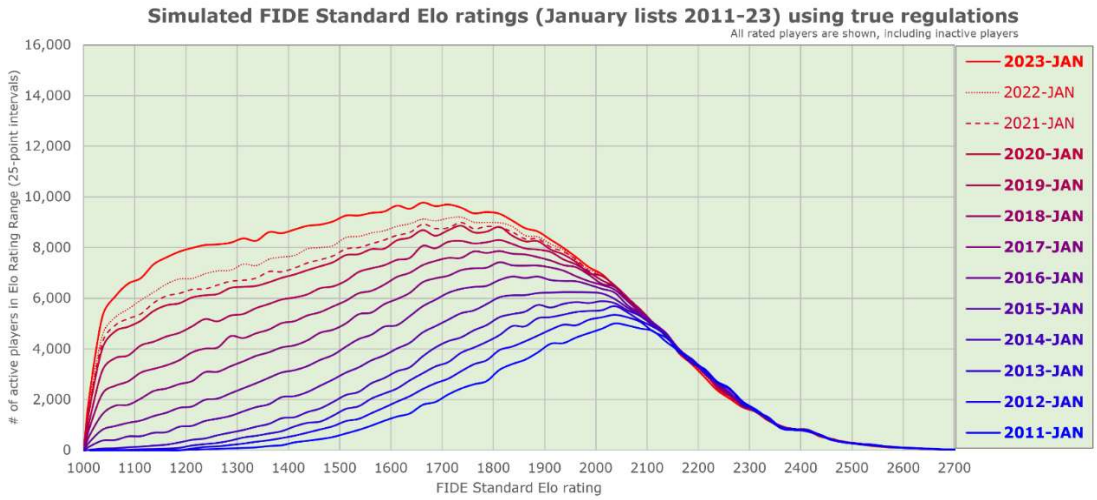
Part V – Simulations

I have tested these various fixes extensively via simulation. The most useful simulation that I ran is probably the scenario where we pretend that both of these fixes were implemented at the start of 2017, meaning that the **Compression** was implemented for the January 2017 rating list and the **Calculation Improvements** were implemented for the February 2017 rating list. This didn't happen in real life, of course, but we can see what would have happened to the ratings if these fixes had indeed happened in 2017 and then all of the millions of games from real life were subsequently played by players within that simulated rating system.

First, we should establish the baseline behavior, namely what happened in real life without these changes. By running the simulator with the actual historical regulations that were in effect from 2008 to 2023, and then counting up how many rated players there were of various ratings on each January list, we can plot them graphically and try to understand how the distribution of rated players has changed over time in the dozen years since 2011.

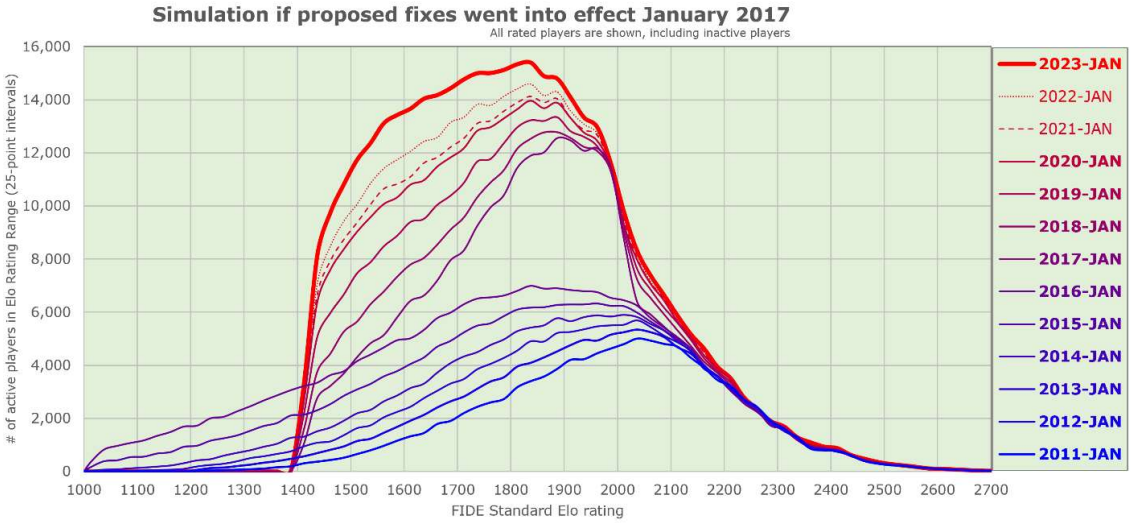
In the following histogram representation, the number of players in each 25-Elo-point range on each January list is counted up, and then plotted on a separate curve for that year. Thus the bottom-most blue curve represents the January 2011 distribution of rated players, the top-most

red curve represents the January 2023 distribution of rated players, and all years in between are also shown, in various red-purple-blue gradient colors.

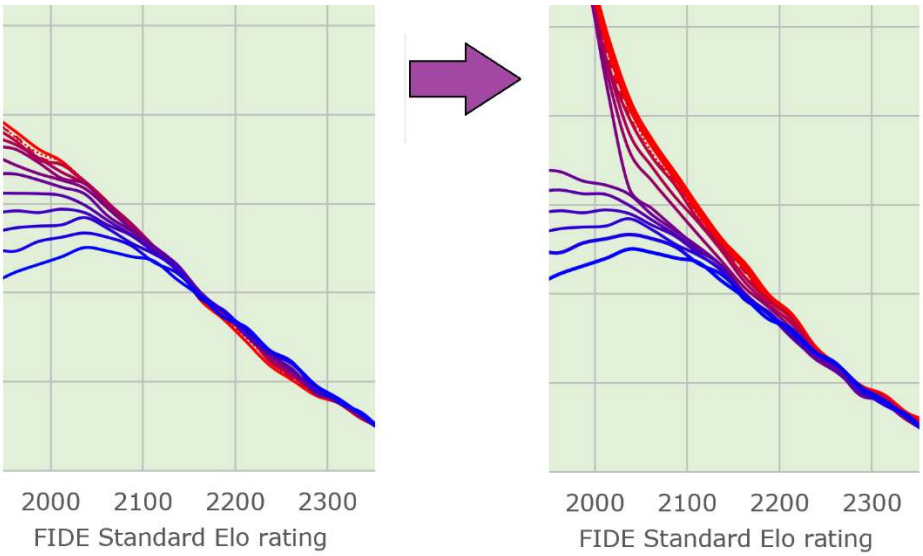


The lines get higher each year on the left half of the image, indicating that there were many more players each year rated below 2100. The rating pool in general grew from about 100,000 players in 2011 to almost 400,000 in 2023. But despite all this growth, we can see from the right half of the image that there's been essentially no growth in the number of players rated 2100+ Elo. This doesn't really make sense at first. Surely with so many new players getting ratings, some of them are master-strength? The explanation, of course, is that whatever modest growth there has been in strong players has been canceled out by the deflation, by master-strength players unfairly losing rating points. Essentially the rating pool has been growing leftward rather than upward.

Under the simulation where we pretend that the **Compression** and the **Calculation Improvements** had been implemented in January/February of 2017, nothing is different on this graph up through 2016, and so the blue lines will look much the same. However, in 2017 things start changing a lot. The bulk of the rating pool gets pushed together by the **Compression**, and then the initial rating formula starts pushing more new ratings toward the middle. The minimum rating stays at 1400 Elo rather than 1000 Elo. So six years later there is a very different-looking distribution of rated players, namely the thick red line indicating the simulated January 2023 rating list:



An obvious feature of the compression is that it leads to many more players rated near 2000 Elo, even if it doesn't immediately produce any additional ratings above 2000 Elo. It is inevitable that some of those players will advance to 2100 Elo or even 2200 Elo. We can once again expect some modest growth in the number of master-level players. Let's zoom in to the region of 2000-2300 Elo from the last two pictures.

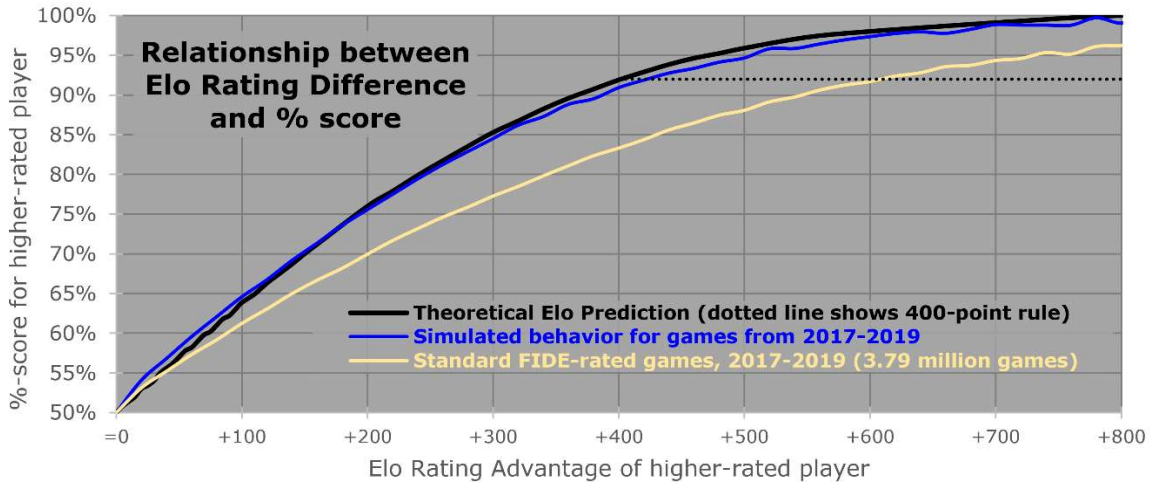


We can see on the left-side image (what really happened) that the red lines didn't get any higher than the blue lines from previous years, meaning there was no growth in this part of the rating pool. However, in the right-side image (the simulation if the compression had happened in 2017), we see from the red lines getting a bit higher each year, that there would once again be some growth in the number of players rated 2100-2200 and a tiny amount of growth in those rated 2200-2300. This is not artificial growth, but a more accurate representation of what's really been happening with the distribution of playing strengths. The **Compression** and the **Calculation Improvements** are an attempt to bring ratings into line with players' evident strengths.

These changes would likely lead to a modest increase in the number of titled players in upcoming years, but these titles would be well-deserved, again because the ratings are (hopefully) being brought more into line with players' demonstrated playing strengths. The intention is to reverse a decade's worth of deflation, rather than artificially increasing anyone's rating or granting undeserved titles. The **Compression** only directly impacts ratings below 2000 Elo, and even players a bit below that would only gain a small number of rating points. My analysis actually pointed to a compression reaching even higher in the rating list, but I felt that stopping it at 2000 Elo would minimize any unwanted effects upon the title system or professional players. The remaining deflation in the system can still be ultimately countered by the **Calculation Improvements**; it just will take longer for this to happen.

We can also use this simulation to determine whether the performance of rated players would be more in line with their Elo expectation (using their simulated ratings, that is). We can take the simulated ratings of all of those players and once again plot them against the black curves indicating their expected percentage score. This will tell us whether rating favorites are actually reaching their Elo expected scores in the simulation.

In the following graphic, the blue curve indicates how rating favorites would have performed in 2017-2019 with these simulated ratings (before the pandemic), whereas the gold curve indicates how rating favorites really did perform in 2017-2019 (before the pandemic) with their actual FIDE-published ratings:



The blue curve almost perfectly overlays/intersects the black curve, which is essentially the goal. This reflects data from simulated ratings during 2017-2019. Even during the simulation, the shutdowns in over-the-board play during the COVID pandemic would have had an adverse effect on the rating system in 2020-2022, causing the ratings of improving juniors to lag behind their true improvement. That is why we are focusing on the performance just in games from 2017-2019 in the above graph, representing the first three years after the **Compression** and the **Calculation Improvements**.

And finally, we can check the summary crosstable from those simulated years of 2017-2019, and see whether rating favorites really are performing closer to their expectation throughout the entire rating pool. As you can see from the following, the red/blue shading is far more pale, and the numbers are much closer to zero, than what happened in real life during 2017-2019. In fact, the numbers look far better than they did even back in 2008-2012.

2017-2019 (simulated)	opponent / player	vs. 1400-99	vs. 1500-99	vs. 1600-99	vs. 1700-99	vs. 1800-99	vs. 1900-99	vs. 2000-99	vs. 2100-99	vs. 2200-99	vs. 2300-99	vs. 2400-99	vs. 2500+
+1.6% in 131,598 games	**1400-99**		**+0%**	**+2%**	**+2%**	**+2%**	**+2%**	**+1%**	**+1%**	**+1%**	**=0%**	**=0%**	**=0%**
+0.8% in 416,053 games	**1500-99**	**-0%**		**-0%**	**+1%**	**+2%**	**+2%**	**+1%**	**+1%**	**+1%**	**+1%**	**+0%**	**=0%**
+0.5% in 612,034 games	**1600-99**	**-2%**	**+0%**		**+0%**	**+1%**	**+1%**	**+1%**	**+2%**	**+2%**	**+1%**	**+1%**	**+1%**
+0.0% in 797,100 games	**1700-99**	**-2%**	**-1%**	**-0%**		**+0%**	**+1%**	**+0%**	**+1%**	**+2%**	**+1%**	**+1%**	**+0%**
-0.5% in 932,399 games	**1800-99**	**-2%**	**-2%**	**-1%**	**-0%**		**-0%**	**-0%**	**+1%**	**+1%**	**+2%**	**+1%**	**+1%**
-0.2% in 957,268 games	**1900-99**	**-2%**	**-2%**	**-1%**	**-1%**	**+0%**		**-1%**	**+0%**	**+1%**	**+2%**	**+2%**	**+0%**
+0.7% in 809,355 games	**2000-99**	**-1%**	**-1%**	**-1%**	**-0%**	**+0%**	**+1%**		**+1%**	**+2%**	**+3%**	**+2%**	**+1%**
+0.2% in 581,912 games	**2100-99**	**-1%**	**-1%**	**-2%**	**-1%**	**-1%**	**-0%**	**-1%**		**+2%**	**+3%**	**+3%**	**+1%**
-0.5% in 418,565 games	**2200-99**	**-1%**	**-1%**	**-2%**	**-2%**	**-1%**	**-1%**	**-2%**	**-2%**		**+2%**	**+3%**	**+2%**
-1.3% in 272,061 games	**2300-99**	**=0%**	**-1%**	**-1%**	**-1%**	**-2%**	**-2%**	**-3%**	**-3%**	**-2%**		**+2%**	**+1%**
-1.5% in 193,010 games	**2400-99**	**=0%**	**-0%**	**-1%**	**-1%**	**-1%**	**-2%**	**-2%**	**-3%**	**-3%**	**-2%**		**+1%**
-1.2% in 123,845 games	**2500+**	**=0%**	**=0%**	**-1%**	**-0%**	**-1%**	**-0%**	**-1%**	**-1%**	**-2%**	**-1%**	**-1%**	

There is every reason to expect that these measures will be very effective at reversing the effects of a decade or more of rating deflation, as well as providing strong counter-measures to prevent it from happening again. This will lead to a FIDE Standard Elo rating system that will be functioning far better than is currently the case.