chess and mind
academia de ajedrez

chess and mind
academia de ajedrez

INICIO   APRENDE   ENTRENA   BLOG   TIENDA

MIEMBROS

MENÚ

Are women worse, or are they scarce? The gender gap in chess – By Alejandro Tello

 BERTRAND_RUSSELL_1903   DICIEMBRE 18, 2020   ANÁLISIS INSTITUCIONAL   0 COMENTARIOS

Are women worse, or are they scarce? The gender gap in chess.

A recent article in Chessbase magazine, written by the professor of psychology and neuroscience and FIDE Master Wei Ji Ma, claimed that the gender gap in top chess is purely explained by population sizes, by the greater participation of men with respect to women.[1] In this article, we will put together all the arguments that have been made with respect to this claim, and we will analyze them for a better understanding of this contemporary issue.

Contents:

1. Introduction to the gender gap in chess.
   a. The gender gap, an apparently obvious fact.
2. Review of the recently published articles on the gender gap.
   a. An attempt to resurrect the participation rate hypothesis – Wei Ji Ma for Chessbase.

PRODUCTOS

Ajedrez: Progresa Aprendiendo Conceptos
★★★★★
20,00€

Ajedrez: Aprende a Pensar Jugando
★★★★★
25,00€

YOUTUBE VIDEOS DE CHESS AND MIND

01) Mamedyarov - ...
00:00   02:01

YOUTUBE SUBS

Chess and Mind en Español
YouTube

COMBINACIONES TÁCTICAS Y SACRIFICIOS

TÁCTICA

Tácticas de Magnus

Jugada tranquila

Mates brillantes

YOUTUBE

Load More...

Subscribe

Review of the recently published articles on the gender gap

## The gender gap, an apparently obvious fact

On the face of it, this theory could be instantaneously dismissed: there is only one woman among the top-200 active players. Only 7 women make the top-500 (Hou, Koneru, Goryakina, Ju, Lagno, Mariya Muzychuk and Anna Muzychuk), while the 10th woman in the ranking (Tan Zhong Yi) is ranked 648th among active players.[2] Not a promising picture for the participation rate hypothesis, considering that about 12% of FIDE affiliations are women.[3] We should keep in mind that, being the distribution of chess ratings approximately normal, the proportion of women should be stable across the distribution, provided that the average and dispersion of the male and female population are the same.

This could be a colossal statistical fluke if it wasn't for the fact that it has been consistent throughout history. Only 2 women have reached the top-50 ever since the first Elo list in 1971: these are the retired chess prodigy Judit Polgar (now insightful chess commentator), and Maya Chiburdanidze, who reached position 48th of the list in June 1988. Along with Hou Yifan, they are the only 3 women who ever made it to the top-100.[4] It seems unlikely that this underrepresentation has occurred by chance or by the sheer differential in participation. Then, how can serious researchers claim otherwise?

Maia Chiburdanidze.

An attempt to resurrect the participation rate hypothesis – Professor Wei Ji Ma for Chessbase
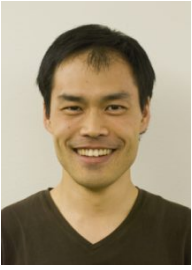
Wei Ji Ma, Professor of Neuroscience and Psychology at New York University. His scientific papers have been cited thousands of times.

Let us analyze what professor Ma did. He took a sample of Indian players from the FIDE rating list. He eliminated junior players from his sample because of their greater rating variability, and found that the average rating of a female players is superior to that of males. But, wait a second, isn't this cherry picking? Junior players constitute more than half the sample after all, and greater statistical variance would average out across such a large population. So I took Professor Ma's sample and verified that his calculations were correct, as it was the case. If the selection of the adult sample was biased, one would expect to get a skewed distribution. However, his distributions were approximately normal (bell-shaped). And in fact his adult male sample had a slightly more positive skewness than the adult female sample.[3] In any case, he used data exclusively from India, and this is indeed cherry picking, because the objective is to analyze the gender gap and not any particular aspect of Indian society. So what about the other countries?

A larger and more diverse sample – Camacho for chess24

A more honest analysis was performed by mathematician and International Master Marcos Camacho, in his article published in chess24 just a few weeks ago. He analyzed the top-10 of each of the top-20 federations in the world according to FIDE. And indeed, men were better in every single country, after correcting for population size, with 16 of those 20 comparisons reaching statistical significance (a standard threshold of confidence used in social science to determine scientific truth). When taking the overall samples (of adult chess populations), 17 of the 20 countries showed average male dominance.[5] Interestingly, 2 countries showed a statistically significant female superiority, one being Georgia and the other India, the very country used in the analysis of the Chessbase article mentioned above.

We must congratulate Professor Ma for finding the one sample in which women are better than men at chess. Apart from this, I see no reason to conclude, as he does with great conviction, that the gender gap in top chess is explained by population sizes.

Federation	Number of players		Gap "top player"		Gap "top 10 players"	
	Men	Women	Expected	Observed	Expected	Observed
RUS	17866 (87.5%)	2559 (12.5%)	72	230*	118	259*
USA	4677 (95.8%)	204 (4.2%)	216	399*	261	357*
CHN	393 (67.9%)	186 (32.1%)	34	133*	61	180*
IND	17899 (93.9%)	1165 (6.1%)	153	167	264	269
UKR	2872 (88.5%)	374 (11.5%)	36	143*	123	245*
ARM	511 (85.0%)	90 (15.0%)	105	360*	197	384*
AZE	614 (82.4%)	131 (17.6%)	106	388*	174	352*
HUN	4865 (93.0%)	364 (7.0%)	148	84	229	234
FRA	17310 (93.4%)	1228 (6.6%)	187	346*	225	349*
POL	7049 (88.8%)	888 (11.2%)	135	342*	153	276*
GER	22595 (94.8%)	1249 (5.2%)	86	200*	163	296*
NED	3938 (95.3%)	193 (4.7%)	206	381*	254	344*
ESP	19895 (96.3%)	774 (3.7%)	190	300*	222	320*
ENG	2828 (95.2%)	142 (4.8%)	185	302	282	345
ISR	1756 (95.4%)	85 (4.6%)	143	349*	260	348*
CZE	6362 (95.7%)	288 (4.3%)	202	305	244	322*
NOR	2667 (96.4%)	99 (3.6%)	413	655*	361	439
CRO	2598 (94.5%)	152 (5.5%)	140	306*	234	372*
GEO	531 (74.1%)	186 (25.9%)	52	162*	75	153*
ARG	3395 (96.5%)	123 (3.5%)	154	284*	252	375*
WORLD	236138 (92.8%)	18401 (7.2%)	91	188*	111	211*

Calculations by José Camacho.

What may explain the biased result of the Indian sample?

Let us now analyze the possible causes for this sharp contrast between the Indian sample and the facts exposed above. As I said, Professor Ma's selection was unbiased. However, if you take the whole of the sample of Indian players (including juniors), males have a very large advantage over female players. On top of that, the ratio of junior to adult players is much larger in the female sample, suggesting that young women could be dropping out in larger numbers than men.[3] I mentioned above that a selection for rating would skew an a priori normal distribution. However, this process of selection is different because it happens progressively. The Elo rating list is constantly updated, and a player's position in the rankings is variable. Therefore, any possible skewness produced by a biased selection would dissipate over time due to a phenomenon called "regression to (or towards) the mean".

Example of a biased selection process

Let us imagine that we select players from the current FIDE list, with a preference for higher ratings. This would make the bottom of the distribution less dense than the top, thus skewing it to the right. However, if we selected young players based on their superior talent, the most talented would sooner or later "regress to the mean" of the rating distribution, and so would the least talented, balancing out the skewness created by the selection process. In addition, the Indian population is highly stratified. Any gender-based selection process of Indian chess players, for instance by social class, would bias the results.

Short summary of the history of this debate in the scientific literature

Old arguments against the hypothesis – Robert Howard

This talent selection process was already pointed out several years ago by Robert Howard. the most prominent detractor of the participation rate hypothesis. He demonstrated that the gender gap increases as the proportion of women increases, and claimed that smaller samples tend to be more selective talent-wise. Countries with greater female participation tend to show a greater gender gap, which is inconsistent with the participation rate hypothesis.[6]
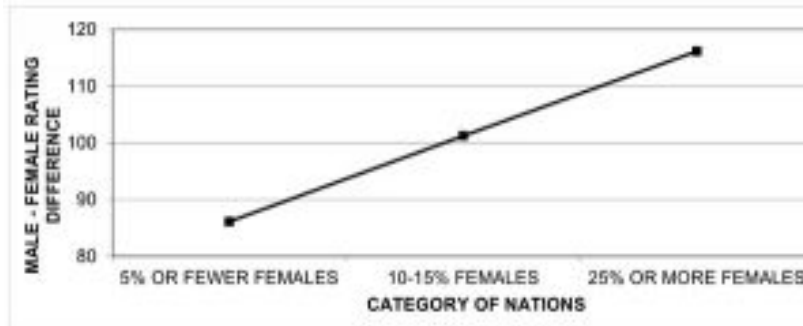
Gender gap over time

Howard also showed that the gender gap has been consistent over time, as I anticipated above. In his paper of 2005, he analyzed data from 1975 to 2004, which he later extended until 2014, showing great consistency at close to 250 points, one standard deviation.[6]
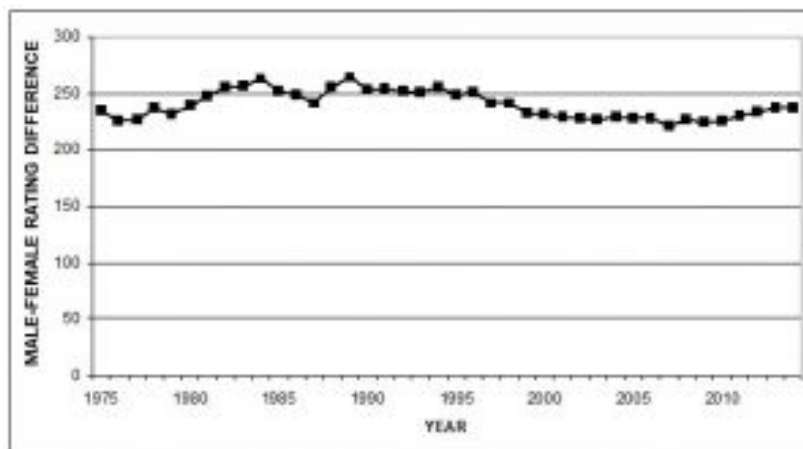
Correcting for confounding factors

On top of this, he adjusted his results to correct for the greater number of games played by the male population. The amount of competitive games played is known to be one of the main predictors of chess strength,[7] so this correction was necessary. He compared women and men with more than 750 games, arguably the number of games necessary to reach one's peak. However, this had no effect and the gender gap remained.[6]
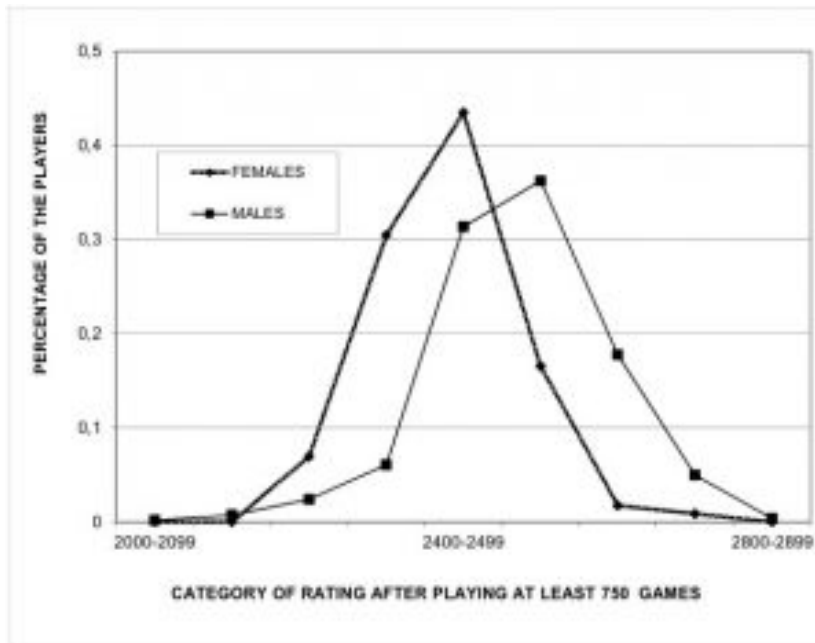
Criticism of Howard's original paper – Merim Bilalic and Peter McLeod 2007

Criticized the original paper by Howard (2005) on the basis that his sample was both corrupted by the lowering of the Elo rating floor by FIDE (the International Federation decreased the minimum rating a player could get), and too selectively based on the top of the distribution. Indeed, FIDE's original rating cutoff was 2200, several standard deviations above the mean of the distribution.[8]

The participation rate hypothesis – Merim Bilalic and colleagues 2009

In a subsequent paper, Kieran Smallbone and Fernand Gobet joined Bilalic and McLeod to analyze a large sample of German players. This sample was not based on FIDE ratings but on a different Elo system by the German Federation, and unlike the FIDE sample it contained information about most tournament players. They created a model to estimate the expected difference of populations of different sizes, and they showed that most of the advantage in rating was due to the greater number of male players.[9]

Refutation of the methods used by Bilalic and colleagues – Michael Knapp 2010

Michael Knapp criticized the method used by Bilalic and colleagues by exposing the flaws of their statistical model. He calculated that, in their own sample, the 100th female player should be ranked between 1510th and 2849th in the overall sample with a degree of confidence of 99.9%, assuming no effect of gender. However, in the actual ranking, the 100th woman was ranked 5505th.[10]
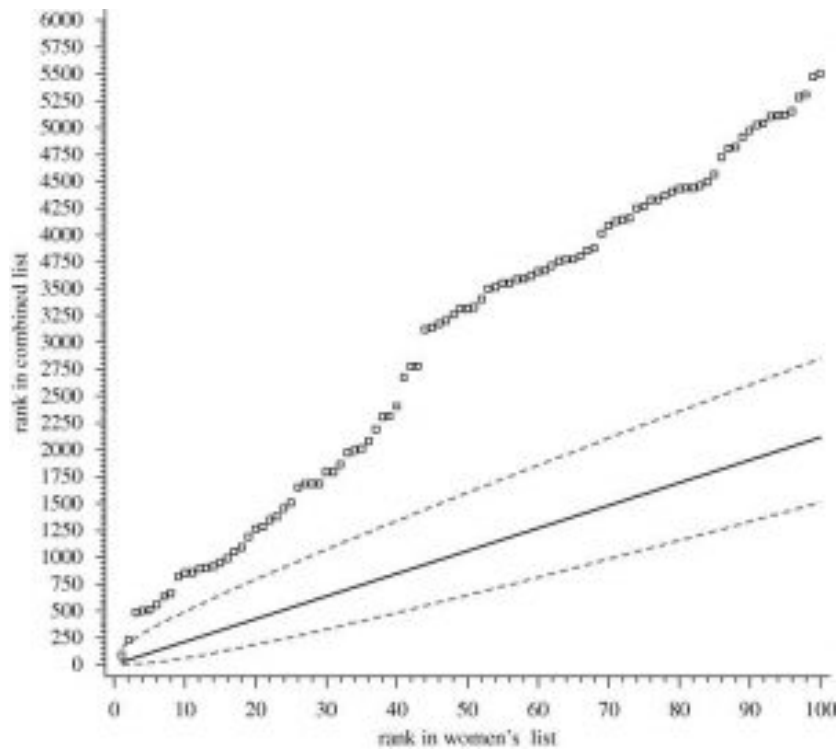
So where did Bilalic and McLeod go wrong? – Bruno Weisend 2019

Apparently, the participation rate hypothesis was born out of a statistical mistake. Bilalic and McLeod assumed that their sample was an approximately normal distribution, with a common mean as a weighted average of the female and male populations. Therefore, their conclusions that the differences at the top were due to population sizes is contingent to this false assumption.[11] We will speak more on this rich and insightful article by Weisend in the future.

Are women worse, or are they scarce? The gender gap in chess - By Alejandro Tello - Chess and Mind

12/24/20, 1:59 PM
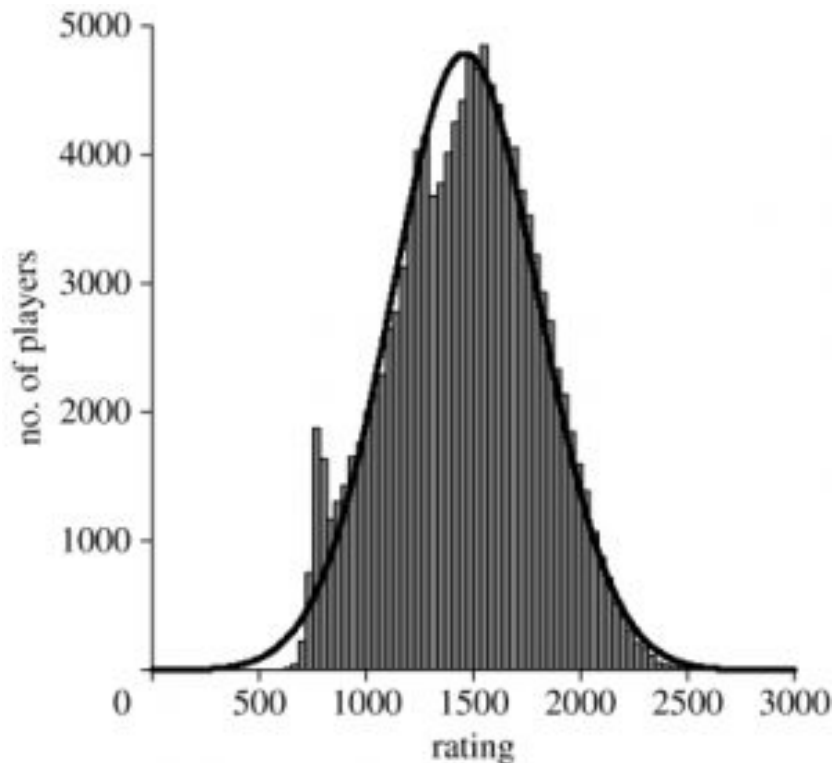
Layman analysis of the problem

Now, let us analyze the issue from a layman point of view. The FIDE Elo ratings are available for everyone to use. Professor Ma calculated the average of female and male Indian players. In effect, when the same procedure applied by Ma is performed for the entirety of the FIDE list, the average of the senior male population is 75.5 points higher than the senior female population.[3] This is over a sample of a quarter of a million people, not precisely a fluke.

Top-countries vs women.

We have already talked of how Camacho found that the vast majority of countries show a statistically significant advantages for men, both in the overall and in the top-player samples.[5] But more impressively, the average gap between the expected and the real difference between the male top-10 and the female top-10 was 98 points! This is, among the 20 countries analyzed by Camacho, the top-10 women were rated almost 100 points lower than expected by the difference in population size. Again, 100 points over such a large sample didn't come by chance.

Country	Expected	Real	Difference
RUS	118	259	141
USA	261	357	96
CHN	61	180	119
IND	264	269	5
UKR	123	245	122
ARM	197	284	87
AZE	174	352	178
HUN	229	234	5
FRA	225	349	124
POL	153	276	123
GER	163	296	133
NED	254	344	90
ESP	222	320	98
ENG	282	345	63
ISR	260	348	88
CZE	244	322	78
NOR	361	439	78
CRO	234	372	138
GEO	75	153	78
ARG	252	375	123
Average	207,6	305,95	98,35
WORLD	111	211	100

Top countries by average top-10.

Small male samples vs women.

One could argue that there is a selection bias in Camacho's sample. Indeed, he took the top-FIDE countries as a sample, consisting on those with the highest top-10 average. Since the male populations are substantially larger, picking the top federations could signify picking the best male populations. In this sense, it could be the case that this gap exists only in the top countries but not in average. For that reason, I made some small corrections to this bias: First of all, we could apply the same procedure but picking countries by female top-10 average instead. As it turns out, the top countries are almost the same whether one picks the female or the male average.[12]

However, a better solution is to apply a completely unbiased method of analysis. If the difference between ratings of women and men are due to population sizes, a subsample of men of the same size as the female population should have a similar top-10. The problem is that there are no countries that even approach the population size of the overall female sample of 145583 (including juniors, unrated and inactive players). The closest country is Russia,[13] with a total male population of 76948 and average active top-10 rating of 2735, while the overall active female top-10 average is 2560.[14]

#	Name	Fed	Rating
1	Hou, Yifan	CHN	2658
2	Goryachkina, Aleksandra	RUS	2592
3	Koneru, Humpy	IND	2586
4	Ju, Wenjun	CHN	2560
5	Lagno, Kateryna	RUS	2546
6	Muzychuk, Mariya	UKR	2544
7	Cmilyte, Viktorija	LTU	2538
8	Muzychuk, Anna	UKR	2535
9	Dzagnidze, Nana	GEO	2524
10	Harika, Dronavalli	IND	2515

Top women's chess ranking October 2020.

100th-matched countries vs women.

There are 23 countries where the average top-10 is higher than the overall female top-10, all of which have chess populations that are overwhelmingly smaller than the female sample. There are some striking cases, such as Armenia, with a male population of only 1503 players (including juniors, non-rated and inactive players), and an average top-10 of 2660, exactly 100 points above the overall female top-10.[14] Similar cases can be made of countries such as Azerbaijan and Georgia.

Extreme examples aside, I took all the countries where the 100th player was close to the overall female of the same ranking, with a rating of 2363. I included countries with a 100th player above 2300: 14 countries in total, all of which have a top-10 average that's higher than the female top-10, with a mean of 88 points. Even when removing the two highest rated federation by top-10 average (Russia and the USA), the difference remained substantial, at 75 points. Suitably, this correction made the average rating of the 100th players in my sample almost identical to the female 100th, 2365 to 2363 respectively. It's also important to notice that 7 of these 12 countries have 100th rankings that are lower than the female 100th, but all 7 have a higher top-10 average.[14]

Ranking	Country	100th player		top-10 average	Difference
1	Rusia	2521		2735	175
2	USA	2400		2713	153
3	China	2310	*	2703	143
4	India	2393		2670	110
5	Ukraine	2405		2665	105
6	Hungary	2350	*	2645	85
7	France	2382		2643	83
8	Poland	2359	*	2634	74
9	Germany	2437		2629	69
10	Netherlanc	2352	*	2628	68
11	Spain	2400		2624	64
12	Israel	2319	*	2610	50
13	Chequia	2327	*	2606	46
14	Serbia	2345	*	2568	8
Average		2379		2648	88
Corrected		2365		2635	76
	Women	2363		2560	-

Countries with the highest rated 100th player.

Adding an extreme outlier to the female ranking.

One could ponder whether this difference is produced by the absence of a strong female player at the top, such as Judit Polgar. First of all, the current female number one, Hou Yifan, is quite unusually strong if we exclude Polgar. In addition to that, eliminating the last top-10 player (Harika, 2515), and adding Magnus Carlsen (2862) to the female top-10 barely changes anything, allowing them to surpass only Serbia and stay behind the other 13 countries. But obviously Magnus is currently rated 125 points above any female player at any point in history

(peaking with Polgar's 2735, July 2005), so this isn't a very plausible scenario.

Simple refutation of the variability hypothesis.

As a footnote, the 100th player is obviously far above of the true average in almost every country, since most countries have more than 200 players. Assuming an approximately normal distributions for male ratings, the average should be close to the median, which is the value that divides the population in half. Of course, ideally we would take the average as a reference point. The problem is that the true average is difficult to calculate, since the FIDE applies a rating floor that wipes out the bottom of the distribution. This is why my argument above only proves that males are higher rated at the top and not in average. It could be that men are overrepresented at both tales, with similar average as women but much greater variability.

Indeed, men seem to show greater standard deviation, but not nearly enough to produce such advantage at the top spots. When removing the upper half of the ranking (of rated players only), both in the male and the female sample, the average rating difference in favor of men goes from 168.6 to 165.1: a very minor change.[3] This means that the gender gap at the top isn't due to greater variability in the male sample; rather, men have a much higher average rating. No permutation test is needed, the p-value is essentially 0. In addition, let's not forget that the FIDE ranking is as close to a census as it gets: most players are rated. There isn't much space for inference errors or for selection biases to occur.

Extracting random samples from the male population

Finally, I did a little more technical and hazardous calculation: I took 100 samples of different sizes from the overall male ranking (of all FIDE affiliations, non-rated players included), and I averaged out the top ten ratings from each sample. I think this is a simpler approach to refute this hypothesis than the method used by Bilalic and others, who tried to extrapolate from a small female sample. It turns out that less than 10,000 male players are needed in order to surpass the female top-10 average (2560). It's important to point out that 100 samples was more than enough, since the standard deviation of these samples was very small (26, with an average of 2575 for the samples of 10,000 men). In fact, the average remains constant from about 30 samples extracted. There are 145,582 women affiliated to FIDE, so random samples of about 1 man for every 15 woman have a higher top-10 average.

Sample size	Average rating	Standard deviation
5000	2523,95	30,94
10000	2575,41	26,12
15000	2603,48	25,67
20000	2620,17	24,59
25000	2635,13	22,63
30000	2644,62	21,35
35000	2652,50	21,37
40000	2660,67	20,30
45000	2666,99	20,06
50000	2673,66	19,55

The average rating refers to the top-10 of the sample.

To put it simply, it makes no sense to say that the differences at the top are due to sample sizes, since a sample of men 15 times smaller than the female chess population has a higher top-10 average. However, this shouldn't be mistaken with the idea that men are 15 times better than women, or that a random man is better than the best of 15 women. It only means that, population sizes are not that relevant for explaining gender differences in top-level chess. It is the greater average rating of males what has the greater weight. It's interesting to note

Are women worse, or are they scarce? The gender gap in chess - By Alejandro Tello - Chess and Mind

12/24/20, 1:59 PM

that a sample of only 50,000 men has an average top-10 that's roughly 120 points lower than the overall male population. In contrast, the female top-10 average is more than 230 points lower, with a population about 3 times larger than the sample mentioned.[3]

In conclusion

The gender gap in top chess is not explained by the difference in population sizes, as professor Wei Ji Ma claims. When matching population sizes, the gap remains substantial.

In this article, I have merely attempted to refute the participation rate hypothesis once and for all. My interest was not to determine how large the gap is exactly or what its causes and mechanisms are. It's well-known that population studies do not show cause and effect. What we can say with great confidence is that men are better than women at chess, in average. Are the differences biological, or are they cultural? Are they due to inherent differences in the wiring of the brains of women and men, as Nigel Short famously argued?[15] Does this mean that men are intellectually superior to or more intelligent than women? These are some of the common questions that will be examined in a future installment.

References

1. Ma,Wei Ji, 2020. What gender gap in chess? [online]
    https://en.chessbase.com/post/what-gender-gap-in-chess

2. Based on FIDE ratings search.
    https://ratings.fide.com/

3. Calculated from the October 2020 FIDE rating list.
    https://ratings.fide.com/download.phtml

4. Extracted from the archives of former FIDE rating lists.
    https://www.olimpbase.org/Elo/

5. Camacho Collados, J., 2020. *The Gender Gap In Top-Level Chess*. [online]
    https://chess24.com/en/read/news/the-gender-gap-in-top-level-chess

6. Howard, R. W. (2013). Explaining male predominance in chess. [online]
    https://en.chessbase.com/post/explaining-male-predominance-in-chess

7. The amount of competitive games played is known to be one of the main predictors of chess strength

8. Bilalic, M., & Mcleod, P. (2007). Participation rates and the difference in performance of women and men in chess. *Journal of biosocial science*, *39*(5), 789.
    https://www.researchgate.net/profile/Merim_Bilalic/publication/6475924_Participation_rates_and_the_difference_in_performance_of_women_and_men_in_chess/links/5d4951cc299bf1995b69838c/Participation-rates-and-the-difference-in-performance-of-women-and-men-in-chess.pdf

9. Bilalić, M., Smallbone, K., McLeod, P., & Gobet, F. (2009). Why are (the best) women so good at chess? Participation rates and gender differences in intellectual domains. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1659), 1161-1165.
    https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2679077/

10. Knapp, M. (2010). Are participation rates sufficient to explain gender differences in chess performance?. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1692), 2269-2270.
    https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2894898/

11. Weisend, B., 2019. Questioning Gender Studies on Chess. [online]
    https://en.chessbase.com/portals/all/2019/05/chess-gender/wiesend_questioning%20gender%20studies%20on%20chess.pdf

12. Top Female rankings by country
    https://ratings.fide.com/topfed.phtml?tops=1
    Top mixed rankings by country https://ratings.fide.com/topfed.phtml?tops=0&ina=1&country

Are women worse, or are they scarce? The gender gap in chess - By Alejandro Tello - Chess and Mind

12/24/20, 1:59 PM

13. Rating analytics: The number of rated chess players goes up.
    https://www.fide.com/news/288

14. (These figures are calculated from FIDE rating advanced searches)
    https://ratings.fide.com/

15. https://www.theguardian.com/world/2015/apr/20/nigel-short-uk-grandmaster-men-hardwired-better-chess-players-women

Related posts:

Is chess24's 'hall of fame' based on 'chessmetrics' a fair ranking? (by Alejandro Tello)

How many female chess players are needed to match the men's top-10?

Share this post

🏷 Etiquetados English

ANTERIOR

Arma letal contra la defensa Pirc | Sorprende a tu riva: aperturas inusuales

SIGUIENTE

How many female chess players are needed to match the men's top-10?

Leave a reply

Default Comments (0) Facebook Comments

Por favor acceder para comentar.

f ▶

About Chess and Mind

Contact

🇪🇸 Español