Understanding the Impact of the Voting Model at the Eurovision Song Contest: Why Croatian Baby Lasagna did not win the ESC 2024?

Nikola Kadoić
University of Zagreb Faculty of organisation and informatics
Department for organisation
Pavlinska 2, HR- 42000 Varaždin, Croatia
nikola.kadoic@foi.unizg.hr

***Abstract**. This paper focuses on the voting system applied in the final of the Eurovision Song Contest (ESC) 2024. The winner was the song that achieved the maximum sum of jury and public points. Each national jury consisted of five music experts. They ranked each song based on various criteria. Then, the national jury points were calculated using the exponential weight model. In this paper, the voting system was demonstrated and critically analysed, and its good and weak points were identified. The ESC24 winner was the representative from Switzerland, who finished only the 5th according to the public votes and the 1st according to the jury votes. Croatian Baby Lasagna was the 3rd, respecting the jury votes, and the 1st, considering the audience votes. The analysis identified at least two unanswered questions. (1) How can the audience be united (by music) when it must agree with their fifth choice? (2) Would national jury points remain the same if another five experts did the jury job? While the first question is more philosophical and not primarily focused in this analysis, answering the second question requires a scientific approach, and the results, if the answer is positive, could help the organisers justify the current jury voting system; on the contrary, if the answer is negative, the organiser has to consider making additional changes to the voting system. The preliminary simulation of differences in national jury points, when four-member juries differ in only one member, indicates the negative answer to the second question. The main conclusion is that the weak points of the voting system are still high, and further changes in the system are mandatory.*

Keywords. Eurosong, public, jury, rankings, votes, music, exponential weight model

1 Introduction

Eurovision Song Contest (ESC) is one of the most popular events in Europe. Its host is the European Broadcasting Union (EBU). Its popularity is also visible in Asia, Africa, and Australia. Many researchers have studied this event in history. While some were more focused on music styles and performance, others were more focused on other aspects, such as the current political environment, marketing strategies, different social issues, inclusion of technology, sentiment analysis for the results prediction, and others. One crucial aspect is the analysis of the voting system, which is the focus of this paper. More precisely, the primary focus is analysing jury voting rules. The main research question of the paper is: *What are the strengths and weaknesses of ESC's current jury voting system?*

At the beginning of this festival, the winner was selected by national juries who gave points to songs and countries. With the development of information and telecommunication technologies (ICT), it was possible to introduce the audience to the voting system. It was and still is not possible to vote for own country. From 1997 to 2008, the winner was determined only by the audience's votes. The juries were just a backup option for the cases experiencing technical problems. A combined approach was applied starting in 2009. The aggregation of national public and jury votes changed throughout the years. At first, the ranks of both components were combined into a single set of points, and later, starting in 2016, each country gave two independent sets of jury and public points (Wikipedia authors, 2024).

The motivation for this paper came from the Croatian result at ESC 2024. Even though Croatian representative Baby Lasagna achieved the best result for Croatia ever (since Croatian independence), 2nd place, many argue about the fairness of the voting system rules, primarily related to the jury part of the voting system. Through the years, the ESC organiser, EBU, recognised the weak points of jury inclusion in selecting winners and consequently constantly applied modifications to the voting system.

Additionally, they recognised the weak points of the only-audience system, which were triggers to return the juries to the game. The official organiser says: "*Using national juries … allows each song to be*

considered individually. It ensures the best qualitative ranking of all participants … "(European Broadcast Union, n.d.). Further, the jury is back to avoid the influence of neighbors and diaspora voting (Michiel Vos, n.d.). However, it is questionable whether this intention has been met and whether the weak points of the only-audience voting system are higher or lower than the combined voting system.

The paper is organised as follows: Section 2 presents the short literature review related to ESC and gives the list of perspectives that were the focus of the interest of different researchers regarding the ESC. Section 3 presents the methodology for the identification of the winner of the ESC. Section 4 presents the exponential weight model applied in the jury voting process. Section 5 presents the identification of the ESC voting system's weak points and good points. Section 6, the concluding section, proposes upgrading the ESC voting systems.

2 A brief literature review on ESC

The idea of this section is not to do a comprehensive literature review related to the ESC but to present examples of research that are related to the ESC and focused on some particular issue. Considering the topic of this paper, the most important is to mention the literature related to the field of Computational Social Choice (COMSOC), which has two heterogeneous layers: (1) Social Choice Theory and related social disciplines (such as Economics, Political science, Philosophy, Psychology) and (2) Computer Science and related disciplines (such as Artificial Intelligence (AI), Mathematics, Data Science, Operational Research) (Dodevska et al., 2020).

The research mainly related to the first layer of COMSOC includes the following research. Cultural diplomacy was the focus of the research by Ginsburgh and Noury (2008). The European identity of ESC was analysed (Carniel, 2015; Coupe & Chaban, 2020). The political situation related to Russian participation in ESC was analysed (Alpatova, 2022). Branding using a cultural perspective was the focus of Baker (2015). Additionally, different biases that influence public voting were analysed by many authors (Budzinski & Pannicke, 2017; Clerides & Stengos, 2006; Dekker, 2007; Dogru, 2013; Mantzaris et al., 2017; Siganos & Tabner, 2020). The media exposure of the countries' representatives was investigated by (Abakoumkin, 2018; Stockemer et al., 2018; Verrier, 2012).

However, the second layer of COMSOC is more interesting for this analysis. The research in this manner is connected to predicting the results, evaluating the voting methods respecting their characteristics, creating new voting methods and participatory models, contest participation, and simulations of crowd voting. The following examples are more oriented to the second layer. Simulations of voting behaviour through the years were analysed in (Budzinski et al., 2023). Further, a simulation of how would the ESC 2021 results look like if the majority judgment method had been applied is presented in (Umair et al., 2022). For purposes of simulating the voting process by crowds, Crowd-PrefRL (Preference-Based Reward Learning from Crowds) can be possibly used, but each user in the crowd must provide a preference label for each preference query (Chhan et al., 2024). A simulation of how social influence affects voting is investigated in (Ganser & Keuschnigg, 2018), and similar simulations can be done in the ESC context.

In the time of wide AI application, it can be used to evaluate alternatives and winner selection in different competitions. The example is given in (Haqbeen et al., 2022). In the case of ESC, a similar application can be made. However, it is questionable whether the organiser and crowd will accept it. Furthermore, a machine learning tool, sentiment analysis, can predict the results. The example related to the ESC 2019 is given in (Kumpulainen et al., 2020).

A question: *Are the jury experts better quality judges than the audience?* was the subject of interest to several researchers in the scientific community, and many individuals without a scientific background were interested in ESC. In terms of scientific research, there are opposite results in different research. In the research conducted in 2005, the authors concluded that jury experts are better quality judges. They came to this conclusion after analysing the results of the ESC before the research was conducted (Haan et al., 2005). Newer research argued, "that televoters are closer to the targets than experts" (Ginsburgh & Moreno-Ternero, 2023). Additionally, some research in progress concluded that the national juries' rankings are less correlated than national audiences' rankings, which is opposite to the expectations—if juries are returned to recognise the quality of songs, how they are so uneven in recognising the quality (they have more different opinions on the quality of song among themselves, than the audience). The research from 2005 could possibly be the EBU's motivation for returning the juries to the competition, but that research should also be observed in light of the period when it was implemented, and the EBU has to rely more on novel research.

Some research proposes new ways of voting and aggregating expert and crowd opinions. A model based on expert and crowd-voting advantages is proposed in (Dodevska et al., 2020). Furthermore, a machine-learning model for the fusion of crowd and experts' opinions was developed in (Kovacevic et al., 2020). The model was also applied to ESC 2016-2018. Another proposal for a framework for the aggregation of crowd and expert opinions based on bargaining theory and optimisation was given in (Vukicevic et al., 2022). Even though the proposed models are substantial and scientifically justified, their main problem is their acceptance by organisers and the crowd. Will somebody accept the voting method if they do not understand it?

3 ESC 2024 (final) voting system

This section presents how the voting system was applied to the ESC 2024 final (European Broadcasting Union, 2024). At the ESC 2024, 37 countries participated in total. Six of them directly went to the finale, and others had to compete in the two semifinals first. In the final, 25 countries competed to win. Ten countries were selected for the final in the first semifinal, and ten were selected in the second semifinal. Five "big countries" participated in the final directly, without competing in the semifinals (Spain, France, United Kingdom, Germany and Italy). Additionally, the winner of the previous ESC, Sweden, went directly to the final. Due to the disqualification of the Netherlands, instead of 26, there were 25 countries in the final.

Even though all 37 countries did not compete in the final, they could participate in the voting procedure of the ESC final. The winner was to be selected using the two sets of points per country: one set is related to the jury points (1-8, 10, and 12), and the other set is related to the public points (1-8, 10, and 12). The process of voting at ESC 2024 is presented in Fig. 1.

3.1 Audience voting

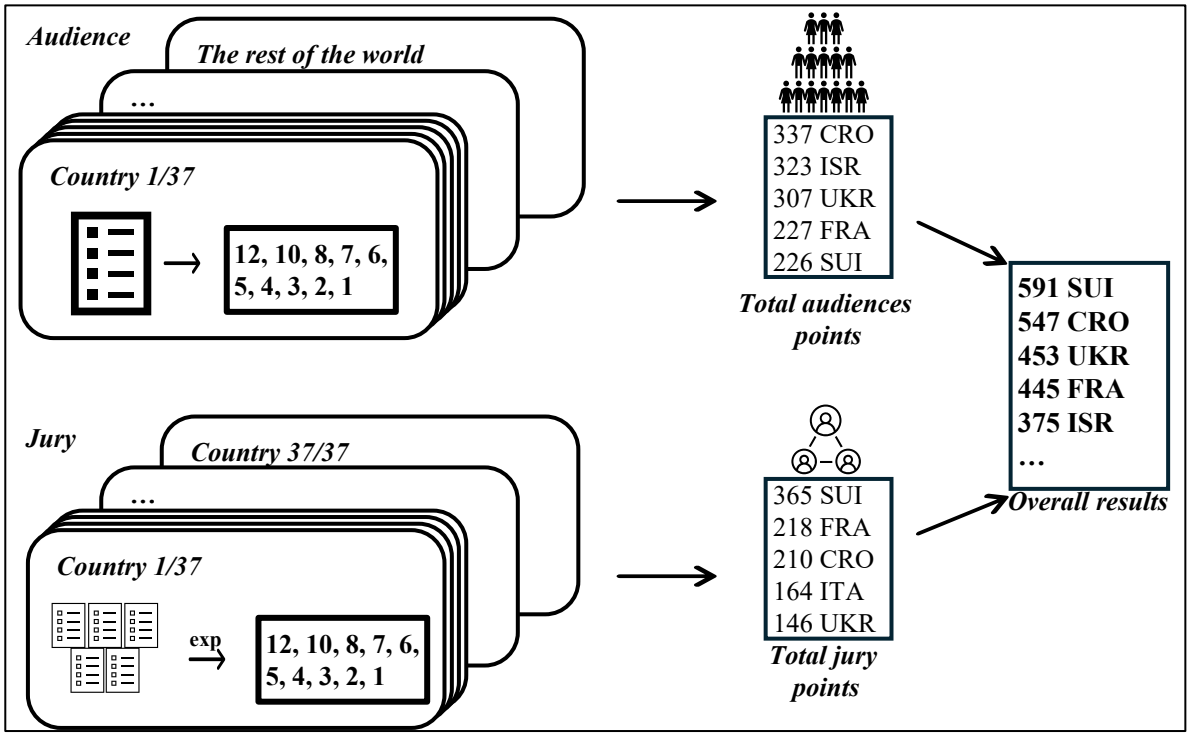
Voters (public, audience) from all 37 countries could vote for their favorites. In each country, after all votes were summed, the first-ranked song received 12 points, the second one received 10 points, the third one received 8 points, and so on, and the tenth-ranked song received 1 point, and all other songs did not receive any points. Additionally, the 38th public—the rest of the world—got one set of points. Consequently, there were 38 sets of public points. In theory, it was possible to achieve up to 444 points from the audience (if a particular country received 12 points from other 36 countries and the rest of the world).

3.2 Jury voting

Regarding the jury voting, the procedure is as follows:

1. Selecting the jury members by national broadcaster. Each of the 37 countries had to create a five-member jury consisting of professionals in the field (members of the national juries must pursue one of the following professions within the music industry: radio DJ, artist, composer, author of lyrics, or music producer). There should also be one backup juror for the case if one of the five members were not to participate in the voting procedure.
2. Familiarising with voting rules. National juries shall vote in all countries and in all cases. All countries shall appoint a national jury to vote in the final (even if their song is not selected for the final). Before voting, the juries had to get familiar with how to vote. The official Public Rules for ESC 2016 states, "*Each national jury shall vote in accordance with the instructions included in the so-called "Green Document"*. However, the *Green document* has never been published. It is not known if the Green Document was still valid in 2024. From what is publicly available, it can be concluded that the document contains instructions on evaluating

Figure 1. The process of voting at ESC 2024



the quality and voting at the ESC. The official ESC website mentioned four criteria that jury members should evaluate: (1) composition and originality of the song, (2) quality of the performance on stage, (3) vocal capacity of the performer(s), and (4) overall impression of the act.
It is not known whether the four criteria have been equally important in making the decisions on voting and whether the songs are evaluated by each juror respecting each criterion separately of having mind on all the criteria in mind at the same time during each song evaluation. Additionally, if the evaluations of songs were implemented using some predefined scale was not known.

3. Preparing for live shows. After the jury members got familiar with their job, they waited for their semifinal and final dress rehearsals. From the official website: "*When voting, Jury Members shall use all their professional skill and experience without favoring any Contestant on the account of their nationality, gender or likeliness and shall be free from bias, external influence foreign perspectives or pressure (including but not limited to betting odds or public opinions). The sole reference for their judgment must be the performance of the Acts during the second (Jury) Dress Rehearsals of the respective Shows."* Some informal ways of communication (such as forums) mention signing statements that confirm that voting will be implemented respecting the previous (and other) instructions.
4. (Voting in the semifinals. Even the focus of this section focuses on voting only in finals, but to get to the bigger picture, here it is noteworthy to mention that the juries also voted in the semifinals (only from competing countries) and that their votes did not affect the semifinals' results but were an option to use jury votes in case of technical problems in some country as a backup option).
5. Voting in the final. Voting in the final was implemented in the dress rehearsal using the instructions jurors received after they agreed to participate in the jury voting (second step). Each juror became the voting form and had to rank all competing songs.
6. Calculating the points respecting the jury votes. After each juror ranked all the competing songs, the ranks were further aggregated into the final national jury points using the exponential weight model (explained in detail in Section 4). There were 37 national juries in the final, and if a country had received 12 points from all 36 remaining national juries, it could have achieved a maximum of 432 points. The distribution of points from the national juries was familiar to the organiser before the final took place because the final dress rehearsal is usually held a day before the final. Some argue that this is a polygon for possible “influences” on the results.

3.3 Results announcement

In the end, after all the songs were performed and after the public voting procedure was finished, it was possible to announce the results.

- Firstly, the jury votes were announced: an announcer from each of the 37 countries was included in the live show announcing the song that received 12 points from their jury, while the other points (1-8 and 10) were presented at the table just before 12-points announcement, during the short, small talk between the announcer and ESC host.
- Secondly, the audience votes were announced in a summative way, starting with the country that was ranked last after the jury votes announcement. So, the hosts read the total number of audience points for the country ranked as last. Further, the total number of audience points were revealed for the country ranked second to last and so on up to the song that is ranked first. This way of revealing the audience's points created an intense atmosphere until the last points were revealed, which was the organiser's goal since the winner was known in some previous ESCs before all the countries presented their votes.

3.4 ESC 2024 specificities

Even though some cheating by juries was not registered, or at least not revealed, in the history of jury voting, there were examples of cheating by the national juries. In 2022, six national juries (Azerbaijan, Georgia, Montenegro, Poland, Romania, and San Marino) were caught cheating. They agreed to vote for each other's. The change from a linear to an exponential weight model in 2018 was also motivated by the cases of favoring or disfavouring some contestants by an individual member of some national jury.

Additionally, each ESC, including ESC 2024, was characterised by different specificities. While some argue that those specificities were just a part of the promoting strategies, others argue that they should not have a place at the ESC. The examples include Israel's participation at the ESC, the exclusion of Russia from the competition, and the scandal with the contestant representing the Netherlands (who made a "threatening move" towards one of the crew members), which was consequently disqualified. All those and other specificities caused some countries to withdraw almost 25 minutes before the final. These events are currently not the focus of this paper, but other researchers will undoubtedly analyse them, focusing on ESC's political perspective. It is possible that those events also had an impact to the final results.

3.5 Open questions

After the last points were announced, it was clear that the winner was Nemo from Switzerland, the leader

after the jury voting. However, he had only the fifth result of audiences. Many argue that the Swiss win is undeserved and that it is not acceptable that the winner is the song, which is the only fifth choice made by the audience. This is not the first time the winner is not the audience's favorite. Similar situations happened before (for example, just one year before, Sweden vs. Finland), but it was the first time that the difference between the winner's rank considering the jury voting (and overall rank) and the winner's rank considering the audience voting was so high.

In this direction, sentiment analysis was conducted among the ESC audience, identified many questions, and some of them were:

1. Is it justified that the juries have an equal set of points as the audience (not considering the rest of the world)? There were 185 jury members from 37 countries in total. On the other hand, there were millions and millions of viewers and voters from all over the world.
2. Is the jury voting system good and fair? Are the ESC voting system's weak points, which include the jury, lower than the weak points of the voting system, which does not include the jury? Has introducing the jury into the game solved weak points of the only-audience voting system or caused new and more significant problems?
3. How do we ensure that juries evaluate the quality of songs objectively? What is the definition of a quality song? What is the measuring instrument related to the quality of songs?
4. Is it justified to promote the ESC slogan *United by music* and at the same time conclude that the winner is only the fifth choice of the audience? How can anything and anyone be united if they have to agree on their fifth choice?
5. Is the Eurovision organised for the audience or juries?

In this paper, the previous questions were not entirely answered but at least partly analysed, and this analysis could be potentially valuable in designing a better voting system, if possible.

4 Exponential weight model vs. linear weight model in jury voting

In this section, a demonstration of the exponential weight model is presented. The exponential weight model was introduced in 2018. Before 2018, the linear weight model was applied.

The main difference between the two models is presented in Table 1. The table contains (possibly) real data from the French jury related to the ESC final

Table 1. The difference between linear and exponential weight model

	Linear weight model								Exponential weight model								DIFF	
	1	**2**	**3**	**4**	**5**	**AV**	**R**	**P**	**1**	**2**	**3**	**4**	**5**	**SUM**	**R**	**P**	**R**	**P**
Malta	17	14	8	22	9	14	15		0,5745	1,0157	3,1748	0,2222	2,6256	7,6128	15		0	0
Albania	6	12	20	14	16	13,6	13		4,642	1,4851	0,3249	1,0157	0,6947	8,1624	13		0	0
Czech Republic	13	3	6	21	11	10,8	8	**3**	1,2281	8,2072	4,642	0,2687	1,7957	16,1417	8	**3**	0	0
Germany	16	23	11	7	22	15,8	19		0,6947	0,1738	1,7957	3,839	0,2222	6,7254	17		2	0
Russia	15	21	14	13	13	15,2	16		0,84	0,2687	1,0157	1,2281	1,2281	4,5806	19		-3	0
Denmark	8	13	9	8	7	9	7	**4**	3,1748	1,2281	2,6256	3,1748	3,839	14,0423	10	**1**	-3	3
San Marino	22	25	25	25	25	24,4	25		0,2222	0,1257	0,1257	0,1257	0,1257	0,725	25		0	0
North Macedonia	3	15	2	15	2	7,4	4	**7**	8,2072	0,84	9,924	0,84	9,924	29,7352	4	**7**	0	0
Sweden	5	7	3	4	3	4,4	2	**10**	5,6131	3,839	8,2072	6,7873	8,2072	32,6538	2	**10**	0	0
Slovenia	25	10	16	24	23	19,6	23		0,1257	2,1714	0,6947	0,152	0,1738	3,3176	22		1	0
Cyprus	9	9	4	17	17	11,2	10	**1**	2,6256	2,6256	6,7873	0,5745	0,5745	13,1875	11		-1	1
Netherlands	2	6	5	2	1	3,2	1	**12**	9,924	4,642	5,6131	9,924	12	42,1031	1	**12**	0	0
Greece	14	22	15	23	15	17,8	22		1,0157	0,2222	0,84	0,1738	0,84	3,0917	23		-1	0
Israel	21	11	18	11	6	13,4	12		0,2687	1,7957	0,4751	1,7957	4,642	8,9772	12		0	0
Norway	23	20	24	18	14	19,8	24		0,1738	0,3249	0,152	0,4751	1,0157	2,1415	24		0	0
United Kingdom	20	18	23	6	20	17,4	21		0,3249	0,4751	0,1738	4,642	0,3249	5,9407	18		3	0
Iceland	10	1	1	20	24	11,2	10	**1**	2,1714	12	12	0,3249	0,152	26,6483	6	**5**	4	-4
Estonia	12	17	19	12	18	15,6	17		1,4851	0,5745	0,3929	1,4851	0,4751	4,4127	20		-3	0
Belarus	24	8	7	19	21	15,8	19		0,152	3,1748	3,839	0,3929	0,2687	7,8274	14		5	0
Azerbaijan	4	4	13	5	4	6	3	**8**	6,7873	6,7873	1,2281	5,6131	6,7873	27,2031	5	**6**	-2	2
Italy	1	2	21	3	12	7,8	5	**6**	12	9,924	0,2687	8,2072	1,4851	31,885	3	**8**	2	-2
Serbia	19	16	17	10	8	14	15		0,3929	0,6947	0,5745	2,1714	3,1748	7,0083	16		-1	0
Switzerland	7	5	12	9	10	8,6	6	**5**	3,839	5,6131	1,4851	2,6256	2,1714	15,7342	9	**2**	-3	3
Australia	18	24	10	1	5	11,6	11		0,4751	0,152	2,1714	12	5,6131	20,4116	7	**4**	4	-4
Spain	11	19	22	16	19	17,4	21		1,7957	0,3929	0,2222	0,6947	0,3929	3,4984	21		0	0
France																		

2019—it was tough to find the actual data (the mysterious *Green document* possibly forbids sharing of them). However, this dataset was found on an informal network (Reddit), and independently of the originality of the data, they will be good enough to demonstrate the exponential and linear weight models. In columns related to the linear weight model, we see how songs were ranked per each of the five French jurors. Ranks are further averaged (AV). Further, the songs were ranked (R) concerning averaged ranks (AV), and finally, the points (P) that would be given to the songs respecting the final ranks (R) if the linear model was used. The exponential weight model assigns fixed values to each position in a rank list, predefined by the EBU to follow an exponential function (e.g., first place = 12 points, second place = 9.924 points, third place = 8.207 points, fourth place = 6.787 points, and so on). Although the exact function is not found in the literature, we approximated it as $14.5591 \cdot 0.52664^{x}$.

In the case presented, ranks from linear and exponential weight models are highly correlated (the Spearman coefficient equals r=0.955). The exponential model gave Iceland and Australia 4 more points than the linear model, while Denmark lost 3 points. The EBU explained that the model change was due to significant ranking differences among jurors. Unlike the linear model, if four jurors rank a song first and one last, the exponential model could still award it 12 points. Conversely, a song ranked last by four jurors and first by one would be rated higher in the exponential model. We conducted a sensitivity analysis using data from Table 1, adjusting some votes. Iceland was ranked first by four jurors and last by one, while San Marino was ranked last by four jurors and first by one. Iceland ranked 1st in the exponential model and 2nd in the linear model; San Marino ranked 25th in the linear model and 11th in the exponential model. Iceland gained 2 points in the exponential model, and San Marino remained at 0 points.

In summary, the exponential model reduces the impact of an outlier juror when their ranking differs from the majority but increases it when they rank a song highly that others do not. However, the impact of "helping" some countries is limited as positions between 25th and 11th score 0 points.

5 Discussion

In this section, the goal is to identify the good sides and the weak points of the ESC voting system. The focus will be more on the weak points related to the jury voting system. The primary goal of the analysis is to emphasise aspects of the voting system that could be considered for improvement.

5.1 Audience voting system analysis

Regarding the audience voting system, some characteristics can be considered good and weak when observed from different perspectives. One is related to using the same set of points (1-8, 10, and 12) for each country regardless of the country size. While one can think that a bigger country has to have more points, others believe that equality in the set of points for all countries contributes to the equal importance of each nation, its culture, and its identity in the European environment. Another characteristic is related to plurality voting (voting only for a particular song(s)). While some can think that it is good to have the opportunity to vote only for some or one song, others think that each televoter should give their opinion on each song. In this direction, Ginsburgh and Moreno-Ternero (2023) endorse the use of the Shapley voting procedure for judges and tele-voters.

One of the weak points of audience voting, which will also be the case in jury voting, is related to the relative grading when distributing the points. As presented in section 3, the song that received the maximum number of votes receives 12 points, the second receives 10 points, and so on. In practice, it is possible that, for example, the first and the second song have tiny differences in terms of achieved votes, while in another case, this difference can be much higher, but in both cases, the points for the first and the second song will be 12 and 10. Probably the fairer system would be to distribute the total number of points (1+2+3+…+8+10+12=58) among the songs, respecting the relative share of votes. For example, this is a general idea in political voting in many countries. This idea will probably not be accepted, and one of the reasons is related to the ESC tradition—the system of points (1-8, 10, and 12) was introduced in 1975, which means that there are almost 50 years of tradition in this particular points-scale usage.

5.2 Jury voting system analysis

The good sides of the ESC jury voting system:

1. The organiser decided to deal with the only-audience voting system's problems, such as diaspora voting and voting for neighbors, by introducing juries that should objectively evaluate the songs and take care of the songs' quality,
2. After returning the juries to the ESC, the organiser is working on tracking the problems of the jury voting system and continuously upgrading the system with modifications that could result in fairer evaluation,
3. Jury members are selected among the experts in the music industry.

On the other hand, the jury voting system has some weak points.

1. The selection of the jury members is challenging, and it is questionable if only five members of the jury are enough for the demanding activity of song evaluation. Additionally, it is questionable whether the jury points from any country remain the same if some members or the complete jury are changed. It will be beneficial for the organiser to order

Table 2. Simulation of research that applies the control groups of jury members

	1	2	3	4	5	SUM-1	SUM-2	SUM-3	SUM-4	SUM-5	R-1	R-2	R-3	R-4	R-5	P-1	P-2	P-3	P-4	P-5	DIFF
Malta	17	14	8	22	9	7,038	6,5971	4,437	7,39055	4,9872	14	14	17	13	16						
Albania	6	12	20	14	16	3,52	6,6773	7,8374	7,14667	7,46769	19	13	12	15	13						
Czech Republic	13	3	6	21	11	14,91	7,9346	11,499	15,873	14,346	8	11	9	7	8	3	0	2	4	3	4
Germany	16	23	11	7	22	6,031	6,5516	4,9296	2,88637	6,50311	16	15	16	21	14						
Russia	15	21	14	13	13	3,741	4,3119	3,5649	3,35248	3,35248	18	19	20	17	20						
Denmark	8	13	9	8	7	10,87	12,814	11,416	10,8675	10,2034	10	8	10	10	11	1	3	1	1	0	2
San Marino	22	25	25	25	25	0,503	0,5993	0,5993	0,5993	0,5993	25	25	25	25	25						
North Macedonia	3	15	2	15	2	21,53	28,895	19,811	28,8951	19,8111	4	2	5	2	6	7	10	6	10	5	5
Sweden	5	7	3	4	3	27,04	28,815	24,446	25,8664	24,4465	2	3	4	4	4	10	8	7	7	7	3
Slovenia	25	10	16	24	23	3,192	1,1461	2,6228	3,16547	3,14368	20	24	22	18	21						
Cyprus	9	9	4	17	17	10,56	10,562	6,4001	12,6129	12,6129	11	9	14	9	10	0	2	0	2	1	2
Netherlands	2	6	5	2	1	32,18	37,461	36,490	32,1792	30,1032	1	1	1	1	2	12	12	12	12	10	2
Greece	14	22	15	23	15	2,076	2,8694	2,2516	2,91782	2,25164	22	22	23	20	23						
Israel	21	11	18	11	6	8,709	7,1816	8,5021	7,18155	4,33522	12	12	11	14	17						
Norway	23	20	24	18	14	1,968	1,8165	1,9891	1,66637	1,12579	23	23	24	23	24						
United Kingdom	20	18	23	6	20	5,616	5,4657	5,7669	1,29873	5,61584	17	17	15	24	15						
Iceland	10	1	1	20	24	24,48	14,648	14,648	26,3233	26,4963	3	7	7	3	3	8	4	4	8	8	4
Estonia	12	17	19	12	18	2,928	3,8381	4,0196	2,92751	3,93747	21	20	18	19	18						
Belarus	24	8	7	19	21	7,675	4,6526	3,9884	7,4345	7,55869	13	18	19	12	12						
Azerbaijan	4	4	13	5	4	20,42	20,416	25,975	21,5901	20,4159	5	5	3	6	5	6	6	8	5	6	3
Italy	1	2	21	3	12	19,88	21,961	31,616	23,6778	30,3999	7	4	2	5	1	4	7	10	6	12	8
Serbia	19	16	17	10	8	6,615	6,3136	6,4337	4,83685	3,83337	15	16	13	16	19						
Switzerland	7	5	12	9	10	11,9	10,121	14,249	13,1085	13,5627	9	10	8	8	9	2	1	3	3	2	2
Australia	18	24	10	1	5	19,94	20,26	18,240	8,41154	14,7984	6	6	6	11	7	5	5	5	0	4	5
Spain	11	19	22	16	19	1,703	3,1055	3,276	2,80374	3,10549	24	21	21	22	22						
France																					

independent research, at least in some countries, that will simulate the voting process using one or more control groups. Achieving the results of this kind of research, that there is no difference in ranks and points between original jury and control juries, will help the organiser to argue the existence of a jury in the ESC voting model. Table 2 describes the idea of the independent research proposal: Let us assume that only four jury members are defined by the Green document (which is very close to five members, which was in reality). Since we have the ranks of all five jury members in Table 1, we can create five four-member juries that will differ in only one member. In Table 2, columns SUM-1 to SUM-5 present the sums of exponential values of five four-members (ex., SUM-1 means the sum of exponential values of a four-member jury consisting of jurors 2-5). Further, columns R-1 to R-5 present the ranks of each of the five four-member juries. Finally, columns P-1 to P-5 present the points for different countries of each of the five four-member juries. Consequently, we see maximal differences (column DIFF) in all countries that receive some points (the differences are between 2 and 8 points).

Especially this is emphasised in the case of North Macedonia, Australia, and Italy. In the cases of North Macedonia and Australia, the final jury points differ by five points. In the case of Italy, the difference is 8 points.

a. Additionally, Spearman rank correlations were calculated to examine the correlations among every two sets of ranks respecting 11 countries receiving the points, and coefficients were between 0.63 and 0.84 which are low values considering the context.
b. Furthermore, we calculated the sums of absolute differences in the number of points between each two juries. The sums vary from 16 (the differences between jury-3 and jury-5) to 26 (the differences between jury-2 and jury-5).

Even though in this simulation, we had one jury member less than in the actual case, when comparing each two juries, they differ in only one member—results indicate that significant differences in the final jury points can appear just in the case of one country—what to expect if, in the control groups, there would be the juries with entirely different members?

2. Not publishing the *Green document* contributes to the dissatisfaction and unacceptance of jury inclusion in the ESC voting system.

3. When evaluating the songs respecting the four criteria, if the jurors have to evaluate each song concerning all criteria simultaneously, then there is a potential problem with the trustworthiness of the decision-making process. The evaluation of the songs, concerning the four criteria, is a multi-criteria decision-making problem. According to Saaty and Ergu (2015), a song evaluation method could be evaluated as *low,* respecting many of the criteria defined for methods evaluation, which does not consequently ensure the used method is suitable for evaluating the song quality.
4. Additionally, when discussing the concept of quality, a complex construct, it is unknown how good, deep, and precise the concept of quality in terms of music is defined in jury members' instructions in the voting process.
5. Like in the case of the audience's voting system, the absolute differences among the songs are not considered on both the level of each juror and the level of the jury. Additionally, if certain jurors find two songs equal in terms of quality, they must give an advantage to one song, which is opposite to the concept of objective song evaluation.
    a. The song that is ranked the first for each juror will receive a value of 12 according to the exponential weight model; the second one will receive a value of 9,924, and so on, even though the absolute difference in quality between any two songs does not have to correspond to the fixed difference of exponential weight model values.
    b. At the level of the whole jury, the points are associated with countries respecting the rank of the sum of exponential values of all five jury members, not based on the relative share of country sum in all countries' sum.

    Consequently, we have a double relativisation of jury voting in the system.
6. While the goal of the exponential weight model is argued, the organiser does not argue with the selection of exponential weight model values for each rank at the level of a juror.
7. The organiser also does not argue that there is an almost equal influence of jury and public votes in the voting system. In the case of ESC 2024, there were 185 judges and millions of public voters.
8. The exponential weight model is introduced to deal with outlier opinions in the jury. The aim is "*to place the judgment of a group of jurors above the opinion of a single individual juror*". Why is someone's right to have the outlier opinion decreased? What if the "problem" is in four other judges?
9. In informal circles, many mention lobbying at the ESC. Since the number of judges is finite and not so high, it is potentially possible to influence their opinion indirectly.
10. It is interesting to analyse the similarity between jury votes among different countries. This analysis is related to the ongoing research currently in the phase of paper review (partly same authors). That research concluded that the public points from different countries are more similar than the jury points. The question that consequently arises is—if the juries are returned to the competition to ensure objectivity, how do they think differently about the quality? On the other hand, if the public cannot evaluate the quality, why do they have better similarity among themselves?

6 Conclusion

The answer to the question from the title is simple—when public and jury points are summed, Croatan Baby Lasagna achieved the second-best score and consequently did not win the ESC 2024. However, the purpose of the title question is not to ask simple questions but to tackle the readers as well in the scientific community (and there are many fans of Baby Lasagna) to critically reason on the voting system characteristics to get united in brainstorming the proposals that can potentially be useful for voting system improvement.

This analysis opened the discussion on several aspects of the ESC voting system, but many aspects were not even touched. However, two aspects are imposed as the most critical questions:

1. How is it possible to create a valid jury (considering the analysis presented in Table 2)? The preliminary simulation shows (too) high differences among four-member juries differing in only one member at the level of only one country.
2. How can millions be united when they must agree with their only fifth option?

Further, many other shortcomings of the jury voting system presented in this research need the organiser's attention. The main conclusion is that the voting system still needs significant upgrades. Maybe the best conclusion summarising the whole analysis is the one by Baby Lasagna: "*The audience will come to my concert, not the jury"*. Indeed, Baby Lasagna sold out many concerts after the ESC 2024, which is not the case with the ESC winner.

There are two main limitations of this research. The first relates to the subjectivity of the paper's author, who is a fan of the Baby Lasagna. However, a great effort is put into objectively arguing and elaborating all the conclusions regarding the good and weak points of the ESC voting system. The second limitation is related to the unavailability of some data related to details of the whole voting system process and motives that support the voting system being designed in the way it is designed. Consequently, some reasoning is based on the data available on different forums and discussion groups, which are not official data and might be incorrect.

Acknowledgment

Thanks to Baby Lasagna for being a motivation for this paper and a role model for many life aspects.

References

Abakoumkin, G. (2018). "Play it, Sam", again and again: Further instances of familiarity effects in the Eurovision Song Contest. *Hellenic Journal of Psychology*, *15*(2), 125–137.

Alpatova, O. E. (2022). 'A Song for Europe': Music or Politics? *Journal of Globalization Studies*, *13*(2), 105–118. https://doi.org/10.30884/jogs/2022.02.07

Baker, C. (2015). Introduction : gender and geopolitics in the Eurovision Song Contest. *Contemporary Southeastern Europe*, *2*(1), 74–93.

Budzinski, O., Gaenssle, S., & Weimar, D. (2023). Disentangling Individual Biases in Jury Voting: An Empirical Analysis of Voting Behavior in the Eurovision Song Contest. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4343866

Budzinski, O., & Pannicke, J. (2017). Culturally biased voting in the Eurovision Song Contest: Do national contests differ? *Journal of Cultural Economics*, *41*(4), 343–378. https://doi.org/10.1007/s10824-016-9277-6

Carniel, J. (2015). Skirting the issue: finding queer and geopolitical belonging at the Eurovision Song Contest. *Contemporary Southeastern Europe*, *2*(1), 136–154.

Chhan, D., Novoseller, E., & Lawhern, V. J. (2024). *Crowd-PrefRL: Preference-Based Reward Learning from Crowds*. https://doi.org/10.48550/arXiv.2401.10941

Clerides, S., & Stengos, T. (2006). Love thy Neighbor, Love thy Kin: Voting Biases in the Eurovision Song Contest. In *https://EconPapers.repec.org/RePEc:ucy:cypeua:1-2006*.

Coupe, T., & Chaban, N. (2020). Creating Europe through culture? The impact of the European Song Contest on European identity. *Empirica*, *47*(4), 885–908. https://doi.org/10.1007/s10663-019-09461-6

Dekker, A. H. (2007). The Eurovision Song Contest as a "Friendship" Network. *Connections*, *27*(3), 53–58.

Dodevska, Z. A., Kovacevic, A., Vukicevic, M., & Delibašić, B. (2020). Two Sides of Collective Decision Making - Votes from Crowd and Knowledge from Experts. *Lecture Notes in Business Information Processing*, *384 LNBIP*, 3–14. https://doi.org/10.1007/978-3-030-46224-6_1

Dogru, B. (2013). Modeling Voting Behavior in the Eurovision Song Contest. In *Munich Personal RePEc Archive*.

European Broadcast Union. (n.d.). *Frequently asked questions #Why allow Juries to vote in the Grand Final if their votes are no longer used in the Semi-Finals?* . Eurovosion.TV . Retrieved April 9, 2024, from https://eurovision.tv/voting-changes-2023-faq

European Broadcasting Union. (2024). *Eurovision Song Contest - Official website*. Eurivision Song Contest .

Ganser, C., & Keuschnigg, M. (2018). SOCIAL INFLUENCE STRENGTHENS CROWD WISDOM under VOTING. *Advances in Complex Systems*, *21*(6–7). https://doi.org/10.1142/S0219525918500133

Ginsburgh, V., & Moreno-Ternero, J. D. (2023). The Eurovision Song Contest: voting rules, biases and rationality. *Journal of Cultural Economics*, *47*(2), 247–277. https://doi.org/10.1007/s10824-022-09456-5

Ginsburgh, V., & Noury, A. G. (2008). The Eurovision Song Contest. Is voting political or cultural? *European Journal of Political Economy*, *24*(1), 41–52. https://doi.org/10.1016/j.ejpoleco.2007.05.004

Haan, M. A., Dijkstra, S. G., & Dijkstra, P. T. (2005). Expert Judgment Versus Public Opinion ? Evidence from the Eurovision Song Contest. *Journal of Cultural Economics*, *29*(1), 59–78. https://doi.org/10.1007/s10824-005-6830-0

HAQBEEN, J., SAHAB, S., & ITO, T. (2022). Evaluating Rank-Coherence among AI-enabled Ranking, Expert Rating and Crowd Voting for Selecting Winners in Incentivized Large-scale Idea Contest for Creative Works. *Proceedings of the Annual Conference of JSAI 200*, 1–8. https://doi.org/https://doi.org/10.11517/pjsai.JSAI2022.0_1S1IS303

Kovacevic, A., Vukicevic, M., Radovanovic, S., & Delibasic, B. (2020). CrEx-Wisdom Framework for Fusion of Crowd and Experts in Crowd Voting Environment – Machine Learning Approach. *Communications in Computer and Information Science*, *1260 CCIS*, 131–144. https://doi.org/10.1007/978-3-030-55814-7_11

Kumpulainen, I., Praks, E., Korhonen, T., Ni, A., Rissanen, V., & Vankka, J. (2020). Predicting Eurovision Song Contest Results Using Sentiment Analysis. *Communications in Computer and Information Science*, *1292 CCIS*, 87–108. https://doi.org/10.1007/978-3-030-59082-6_7

Mantzaris, A. V, Rein, S. R., & Hopkins, A. D. (2017). *Examining collusion and voting biases between countries during the Eurovision song contest since 1957*. https://doi.org/10.18564/jasss.3580

Michiel Vos. (n.d.). *Eurovision: Does the voting system with juries need to change?* A Bit of Pop Music. Retrieved April 9, 2024, from https://abitofpopmusic.com/2019/05/24/eurovision-does-the-voting-system-with-juries-need-to-change/

Saaty, T. L., & Ergu, D. (2015). When is a Decision-Making Method Trustworthy? Criteria for Evaluating Multi-Criteria Decision-Making

Methods. *International Journal of Information Technology & Decision Making*, *14*(06), 1171–1187. https://doi.org/10.1142/S021962201550025X

Siganos, A., & Tabner, I. T. (2020). Capturing the role of societal affinity in cross-border mergers with the Eurovision Song Contest. *Journal of International Business Studies*, *51*(2), 263–273. https://doi.org/10.1057/s41267-019-00271-3

Stockemer, D., Blais, A., Kostelka, F., & Chhim, C. (2018). Voting in the Eurovision Song Contest. *Politics*, *38*(4), 428–442. https://doi.org/10.1177/0263395717737887

Umair, A., Masciari, E., Madeo, G., & Habib Ullah, M. (2022). Applications of Majority Judgement for Winner Selection in Eurovision Song Contest. *ACM International Conference Proceeding Series*, 113–119. https://doi.org/10.1145/3548785.3548791

Verrier, D. B. (2012). Evidence for the influence of the mere-exposure effect on voting in the Eurovision Song Contest. *Judgment and Decision Making*, *7*(5), 639–643. https://doi.org/10.1017/S1930297500006355

Vukicevic, A., Vukicevic, M., Radovanovic, S., & Delibasic, B. (2022). BargCrEx: A System for Bargaining Based Aggregation of Crowd and Expert Opinions in Crowdsourcing. *Group Decision and Negotiation*, *31*(4), 789–818. https://doi.org/10.1007/s10726-022-09783-0

Wikipedia authors. (2024). *Voting at Eurovision Song contest*. Wikipedia. https://en.wikipedia.org/wiki/Voting_at_the_Eurovision_Song_Contest