中学校における特別活動（学校行事）としての合唱活動と、その競技性5
― 合唱活動に絶対評価、相対評価を導入するための基礎研究 ―

柴田　篤志

Choral activity, as special activity (school events), and its competitive nature 5: Basic research to introduce absolute evaluation and relative evaluation into choral activities

構成

0.音楽活動を採点することの方法論、絶対評価と相対評価について
　　0-1　前提と仮説
1. 模擬合唱コンクール
　　1-1　課題曲及び自由曲について
　　1-2　採点方法について
2. 模擬合唱コンクール結果
　　2-1　採点結果（絶対評価）
　　2-2　採点結果（相対評価）
3. 考察
　　3-1　絶対評価と標準偏差
　　3-2　絶対評価と相対評価とに生ずる順位の相違、揺れ
　　3-3　絶対評価の妥当性
4. まとめ

0. 音楽活動を採点することの方法論、絶対評価と相対評価について

0-1　前提と仮説

本論は、継続研究の第5稿に当たる。

2021の同タイトル研究（4）では合唱コンクールの採点方法として絶対評価と相対評価のどちらがふさわしいかを論じたが、そこで得られた当面の結論は

→相対評価は順位決定に、絶対評価はフォローアップに効能あり

→データのばらつきの大きさは、絶対評価では起こりがちだが、相対評価で起こることは稀であり、両者が連動している時は表彰の対象とできる可能性あり

→絶対評価と相対評価の標準偏差がともに大きい場合、問題は審査される側にあるとは限らない

であった。同時に「まとめ」において「絶対評価と相対評価では順位に用いる妥当性が変わるのか、という部分については、十分な考察には至らなかった。」としており、前稿（第4稿）では絶対評価・相対評価のどちらが合唱活動（を含む音楽演奏）の評価にふさわしいかの結論を得られていない。

そのため「『相対評価一本』での順位算出は、突き詰めれば人気投票に過ぎない。その根拠がどこにあるかを、絶対評価との併記で可能にできるかも知れない」との視座は本研究で再び仮説として運用するべきものとなった。

本研究は第4稿まとめの末尾に記した、「次回は評価者の数を増やし、その信頼性をあえて落として棄却データについて考える、もしくは、6段階評価が実質的に5段階、4段階となることをなくすため、『持ち点制』評価を試みる」をそのまま使用することを目指す。

本研究では、評価者の数を増やすため、チームあたりの演奏者数を減らしている。

第4稿では全部で53名。参加グループは五つ(12名、8名、9名、12名、12名)であった。採点結果を提出したのは51名だったので、12名グループの採点結果は39名が評価したことになる。延べ評価者数は39、45、42、39、39で204名であった。

本研究は、参加者46名。参加チームは六つ（7名、7名、8名、8名、8名、8名）で、採点結果は全46名が提出。延べ評価者数は39、39、38、38、38、38で230名となり、第4稿に比べて10％ほど増えている。

棄却データを作るほどのデータ増加は叶わなかったが、微増とは言えデータ数が増えたことにより分散の傾向に信頼性が増したと考えたい。また、標準偏差拡大を狙った「持ち点制度」についても試験的に運用している。

さて、研究の内容に踏み込む前に、前提となる「音楽活動を採点（評価）すること」について記しておきたい。学校教育における音楽科では、成績を付す、という形で評価は必ず行っているが、それは必ずしも採点することではない。また、評価とは「優劣を示す」もののみを意味しない。このあたりは学校教育に携わるものには自明のことであっても、特に“採点される側”、児童・生徒にとってはっきりさせておかなければならない問題だろう。

学校において行われる合唱活動は、多くの場合競技性を持っており、順位が付される。そこでは通常採点が行われる、という認識が本研究のベースにあるのだが、教育活動として行われるコンクールであるならその『点』は「多く取ったものが、少ないものに優って

いる」ことを示すとは限らないはずである。だからこそ、「絶対評価」「項目別評価」というものの有用性が強調されるべきであり、教育的効果のみを狙うのであれば“歌唱賞”“調和賞”“言霊賞”“活力賞”などの部門表彰だけ行って「総合評価」としての順位は出さない、というやり方も検討されて然るべき、と考える。

とは言え、参加側には「相対的な優劣」がやはり“わかりやすい”ご褒美になる。だからこその相対評価になるのだが、第4稿まとめに書いた通り、それでは「人気投票になる」可能性も高い。何より、学校行事として行われるコンクールの場合、審査側が音楽的経験にあふれているとは限らず、そうした優れた審美眼を持つ審査員が居たとしてもその数は多くないと考えるべきだろう。いわば「音楽的には素人」が集まって審査したときにその評価にどの程度の信頼性があるのか、もしくはないのかを判断するための基準が必要となるはずである。

本研究の目指すものは、そこにある。絶対評価、相対評価、どちらが信頼に足る評価なのか、その信頼性を増すにはどのような取り組みが必要なのか。特に本稿に於いては、第4稿での「順位決定には相対評価に分がある」をそのまま流用し、「フォローアップには絶対評価が有用」という部分を細く検証していきたい。

1.模擬合唱コンクール

1-1　課題曲及び自由曲について

本稿において、「曲の選定」は考察対象としていない。課題曲はこちらから与えている。自由曲は43曲の候補曲をリストとして渡し、そこから選択させている。第4稿では23曲であったが、単純計算で20曲（実数で32曲）増加させた。これは、教職実践演習の履修者46人を6チームに分けたため、曲の難度の幅を広くする必要があったこと、昨年度のリストから大学生にとって選択する魅力に欠けると判断される12曲を削除したことに伴う措置である。

《課題曲と自由曲》…「曲名」作曲者

◆課題曲

「ほらね、」松下耕

◆自由曲リスト

「IN TERRA PAX」荻久保和明

「omnibus star 光年の旅」鹿谷美緒子

「初心のうた」信長貴富

「旅立ちの時～asian dream song」久石譲

「春に」木下牧子

「ヒカリ」松下耕

「ふるさと（嵐）」youth case
「君とみた海」若松歓
「信じる」（混声版・女声版）松下耕
「名付けられた葉」飯沼信義
「明日へ」富岡博士
「あなたへ　旅立ちに寄せるメッセージ」筒井雅子
「あすという日が」八木澤教司
「YELL」水野良樹
「COSMOS」ミマス
「HEIWAの鐘」仲里幸広
「Believe」杉本竜一
「はばたこう明日へ」松井孝夫
「ぜんぶ」相澤直人
「サッカーによせて」木下牧子
「この地球のどこかで」若松歓
「この星に生まれて」杉本竜一
「かえられないもの」若松歓
「いつまでも」若松歓
「手紙」アンジェラ・アキ
「群青」信長貴富
「空駆ける天馬」黒澤吉徳
「輝くために」若松歓
「川」石桁冬樹
「青い鳥」北山陽一
「心の瞳」三木たかし
「証」阪井一生
「生命が羽ばたくとき」西澤健治
「心のキャッチボール」松長誠
「明日の空へ」山崎朋子
「大地讃頌」佐藤真
「走る川」黒澤吉徳
「夢みたものは」木下牧子
「虹」御徒町凧
「大切なもの」山崎朋子

「僕が守る」（女声版）上田真樹

「友～旅立ちの時～」北川悠仁

なお、前稿でのリストから削除したのは「unlimited」若松歓、「青葉の歌」熊谷賢一、「聞こえる」新美徳英、「今日は君のbirthday」若松歓、「決意」鈴木憲夫、「そのひとがうたうとき」木下牧子、「地球の詩」三浦真理、「ひとつの朝」平吉毅州、「ぼくはぼく」三宅悠太、「予感」大熊崇子、「若い翼は」平吉毅州、「翔る川よ」瑞木薫、の12曲。

六つのチームはこ、ね、め、や、い、もと略記する。自由曲は以下のように決定した。

こ　僕が守る（女声版）

ね　いつまでも

め　虹

や　証

い　友～旅立ちの時～（女声版）

も　あすという日が（女声版）

チーム別の男女構成は男声を含むのがね、め、やの三チーム。それぞれ二人、二人、三人が男性。残りのや、い、もは女声のみである。チーム人数はこ、ねの二チームが七人、その他は八人である。

例年、一チームは八人から十数人で構成したが、「評価者の数を増やす」という方針にそって少人数でのチーム構成となった。その分、声量がなくとも合唱として成立しやすい曲をリストに多く含めたが、六チーム中四チームがNHK音楽コンクールの課題曲（うち一曲は高校課題曲）を選択、いい意味で想定を裏切られた。

1-2　採点方法について

絶対評価に用いる四観点を声量、ハーモニー、表現の工夫、歌詞の理解としている。

第4稿では声量ではなく「音程」を用い、ハーモニーは「ハーモニー･バランス」であったが変更した。まず、音程の場合、減点方式でしか採点されず、音大生の音楽能力で考えて低い評価点は出にくい。ただし、音量であれば声を出す専攻（声楽、ミュージカルなど）は当然アドバンテージがあるため、「大きな声」と「そうでもない声」の差はつく、と考えた。また、大きな声を出せるものが普通に発声してしまうと、ボイスアンサンブルの経験が低い構成員の声が聞こえずハーモニーが損なわれる。このバランス調整を「声量とハーモニー」という二つの項目で評価できるかを問う意味で、今までハーモニーの項目に併記していた「バランス」項目を削除している。こちらの意図として、項目数は四項目で第4稿と同じだが、審査する側に高い音楽理解力を求めている。

採点に関しては、参加六チームの演奏がすべて録画・録音終了し、採点のために授業で使用するサーバーにアップされた段階で次の連絡文を示している。

《連絡文》

手順

模擬合唱コンクール、六チームの録音が出揃いましたのでアップします。

自分のチーム以外の五チームを採点してください。

採点は自由曲、課題曲それぞれについて

「声量」

「ハーモニー」

「表現の工夫」

「歌詞の理解」

の四項目をそれぞれ六点満点（最高が六点、最低が一点）で。全部満点なら24点。

…こちらは絶対評価です。

何を以て最高点とするかは採点者ごとに異なっていてもいいですが、採点者の中では一貫した判断基準を求めます。

これとは別に、（中略）「持ち点割り振り」による順位点を出してもらいます。

こちらは相対評価です。

持ち点は一人15点、五チーム均等に割り振れば三点ずつになります。みんな同じくらい上手で差がつけられない、という場合は全チーム三点、でもいいです。

でも、いいな、と思うチームに四点、五点を出せばどこかのチームは二点、一点になります。

一位から五位までをきちんと点数で区別するなら、五、四、三、二、一点になり合計は15点です。

上位二チームは差がなし、と考えるなら、四、四、三、三、一や、四、四、三、二、二になりますし、一チームだけ図抜けていてほかは…という場合、七、二、二、二、二のような極端な配点もできます。あくまで、他のチームと比べてどの程度優れているかを15点の割り振りで示してください。

「15点を使い切る」こと。余らせてはいけませんし、当然15を超えてもいけません。

絶対評価に関する採点方法は第4稿と同じ六段階評価。こちらの狙いとしては〈普通より良い〉が四、〈普通より劣る〉が三になるように配点することを狙っているが、今まで同様の評価方法を用いた調査ではほぼ二、一は使われず、実質的に四段階評価になっている。今回の調査ではこの絶対評価部分には大きな変更を加えていない。評価者の数が増え

ることだけが前回との違いとなる。

相対評価部分に関する採点方法は大胆に変更した。一人の評価者は、自分の属するチームは採点しないので五チームを採点することになる。これを「一位、二位…」と記述させた場合、絶対評価において同点だったチームを同じ順位にできなくなる。更に、評価者においては絶対評価の項目別評価の合計によって相対順位を出すとは限らない（ある項目のみを優位に判定する）ことを想定し、評価者一人に持ち点15を与え、それを必ず使い切る、という方法を採用した。可能性としては全てのチームが3、全部同じ順位で差はなし、という判定も可能になった一方、突出して一チームだけが優れ、残りはどんぐりの背比べ、という判定も可能になっている。これは、一評価者の配点を標準偏差によって分析することを企図してのものである。

第4稿においては、相対評価は絶対評価に使用した四つの項目を「項目別に」順位だけ出すこと、としていたが、今回は「総合の順位を」「持ち点の割り振りで」順位づけすること、と変えている。前稿において、次回の研究では持ち点制を、としたのは絶対評価の分散を大きくしたかったためであるが、絶対評価と相対評価のどちらが音楽活動の採点評価として望ましいかという「違い」を際立たせるため、今回は相対評価の配点だけを変更している。

2. 模擬合唱コンクール結果

教職実践演習A、B履修者によって1/11、1/13の授業時間に演奏を行った。各チーム持ち時間は15分。所属チーム以外の五チームの演奏を採点、評価するために、すべての演奏は録画し、録音した。なお、昨年度までは講義に用いる教室で、授業録画用のビデオカメラでの録画であったが、今回は音響スタジオの使用許可を申請し、録音技師立ち会いのもと二日間かけて録音している。

ところがあまりに条件が良くなったため、バランスや響きなどが「最高の状態」に整えられた音源となり、これでは審査に差し障るという判断から、スタジオ内で録画用に収録したビデオカメラの音源を「審査用」として供与している。結果的にライブ演奏とほぼ等しい音響データとなった。なお、高音質データは審査集計後、自由に視聴できる形にして学内サーバーで公開している。

2-1　採点結果（絶対評価）

絶対評価四項目の、六チーム分の集計を以下に示す。

	こ				ね			
名前	声量	ハーモニー	表現の工夫	歌詞の理解	声量	ハーモニー	表現の工夫	歌詞の理解
sum	219	197	193	186	185	188	169	162
avr	5.6154	5.0513	4.9487	4.7692	4.7436	4.8205	4.3333	4.1538
標準偏差	0.6331	1.0500	0.8255	1.0121	0.8497	0.9140	0.8686	0.8124

	め				や			
名前	声量	ハーモニー	表現の工夫	歌詞の理解	声量	ハーモニー	表現の工夫	歌詞の理解
sum	183	147	139	139	160	126	135	133
avr	4.8158	3.8684	3.6579	3.6579	4.2105	3.3158	3.5526	3.5000
標準偏差	0.8336	0.9349	1.0208	0.9939	0.9907	1.1879	1.0319	0.9227

	い				も			
名前	声量	ハーモニー	表現の工夫	歌詞の理解	声量	ハーモニー	表現の工夫	歌詞の理解
sum	168	168	170	161	157	169	137	133
avr	4.4211	4.4211	4.4737	4.2368	4.1316	4.4474	3.6053	3.5000
標準偏差	0.9482	0.9482	1.0587	1.1255	0.9911	1.2236	0.9737	1.0591

項目別に順位を出すならば、平均値に基づくことになる（チーム構成員の数が異なるので小計では評価者の数が異なる）。

声量　　こ→め→ね→い→や→も

ハーモニー　　こ→ね→も→い→め→や

表現の工夫　　こ→い→ね→め→も→や

歌詞の理解　　こ→い→ね→め→やも（同値）

こは四項目全てで一位、ねは一項目で二位、三項目で三位。この二チームが総合評価すれば一位、二位になる。

その他はいが二項目で二位、二項目で四位。めが一項目で二位、二項目で四位、一項目で五位。もが一項目で三位、二項目で五位（歌詞理解は同順位二チーム）、一項目で六位。やが二項目で五位（歌詞理解は同順位二チーム）、二項目で六位。

各項目の順位を見た限り、こ→ね→い→め→も→やという順位になると予想できる。

複数項目の絶対評価を如何に順位に結びつけるかは本研究の検証課題からははずれるが、仮に「一位六点、二位五点…六位一点」と点数化すればこ24、ね17、い16、め13、も8.5、や5.5となり、予想された順位と等しくなる（同点五位のやもは1.5点とする）。

こんな強引な数値化を行わずとも、項目平均値を見ただけでこは五点台が二項目、四点代後半が二項目、と圧倒的に高い評定値を獲得している。ねは全て四点台、うち二項目は

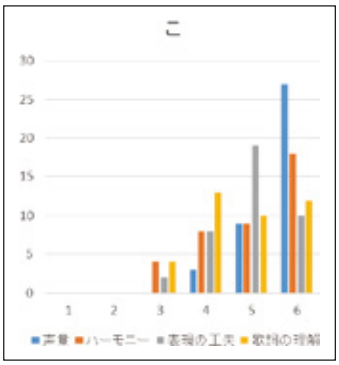
四点代後半、やはり抜きん出ている。この両チームの差は、各項目平均値において明らかであり、一位二位はこれだけで決まるとすら考えられる。翻って、三位以下の順位付けは果たして絶対評価での評定値に妥当性を認めるべきか、まだ結論はできない。

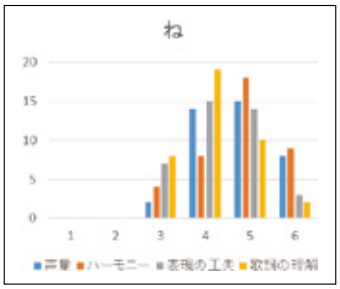
ねといは全て四点台、「表現の工夫」「歌詞の理解」ではいがねを上回っており、順位の逆転現象が見える。偶々、「仮の順位点」による順位付けではねが二位（17ポイント）、いが三位（16ポイント）となったが、本来この両チームには明確な評定上の差異はないと考える。

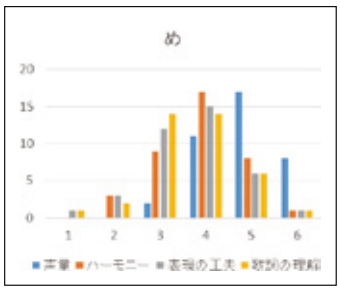
大きなファクターとなるのは標準偏差で、ねの標準偏差は全て1以内に収まるのに対し、いの標準偏差は、いがねを上回る二項目において1を上回っている。これは〈評価が割れて〉結果として高い評価をえたということで、「非常に良い」とした評価者の出した評定が平均値に大きく影響したことになる。

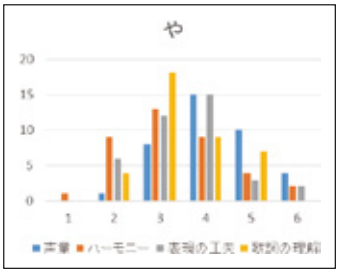
いとめとの差は大きいので三位、四位の順位決定は問題ないとして、四位と五位、めともとの差は明白とはいえない。めともの場合、めの上回る三項目は僅差であるのに対し、もが優る一項目（「ハーモニー」）は圧倒的に上回っている。ものハーモニーの評定地は、三位いをも上回っている。そして、やはり見て取れるのは標準偏差であり、ものハーモニーは標準偏差が1.2236で全チーム全項目の中で最大の値である。評価が割れた結果の高評価、ということになる。複数の評価者で絶対評価を行った場合、どの評価者の行う評価が妥当であるかの判断ができない場合、「データの棄却」によって“妥当な評定値”に数値を修正しなければならないが、その必要性がよく見える典型だと考える。

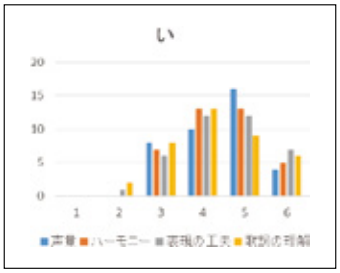
各チームの四項目評価を以下にグラフで示す。それぞれのチームが項目毎に1～6の評定値をどのくらいの数獲得したかを棒グラフの色で表している。なお、評価者の数はこ、ねが39名、い、め、も、やが38名である。

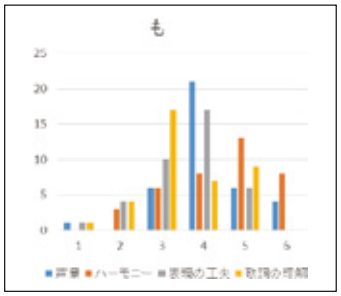












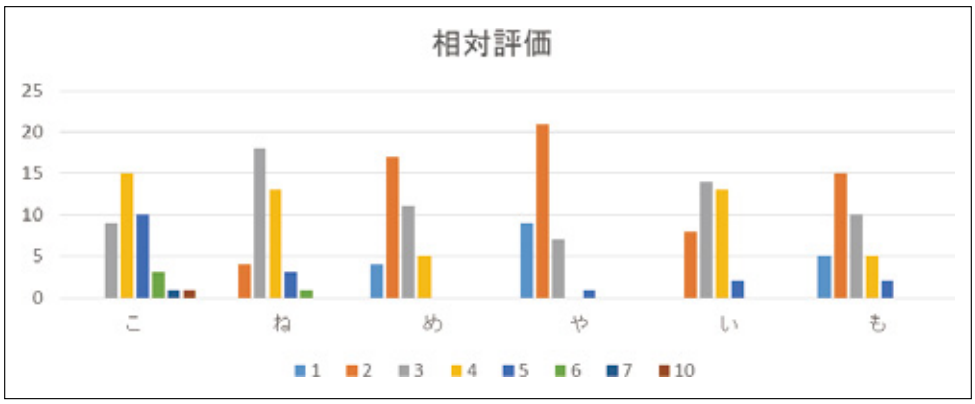
2-2　採点結果（相対評価）

相対評価の、六チーム分の集計を以下に示す。各々の評価者の出した評定値をそのまま単純に合計し、評価者数で割って平均を算出している。

名前	こ	ね	め	や	い	も
sum	172	135	91	77	120	95
avr	4.4103	3.4615	2.3947	2.0263	3.1579	2.5000
標準偏差	1.3518	0.8840	0.9455	1.0761	1.0007	1.1330

相加平均を基にして順位を出せば、こ→ね→い→も→め→やとなる。絶対評価の「仮数値を用いた四項目の総合順位」がこ→ね→い→め→も→やであった。四位と五位が入れ替わっている。これに関しては絶対評価の項目で既に触れているが、「絶対評価における標準偏差」に関する問題として後述する。

標準偏差は相対評価でも算出しているが、こちらの分散は絶対評価よりもかなり大きくなる。もっとも大きな分散を示しているのは一位となったこであり、獲得した評定値の値自体は突出していても、高い評価、低い評価が混在、散らばっていることがわかる。逆の意味で注目出来るのが絶対評価でも相対評価でも等しく「二位」となったねで、標準偏差が0.8840と最小になっている。この六チームを並べて優劣を決めたとき、順位としてのランクと「総意として評価がまとまっている（標準偏差が小さい）こと」が相関していない。



なお、各チームの獲得点分布はグラフにして示す。

3. 考察

3-1　絶対評価と標準偏差

2-1で言及した、ねといの順位逆転を標準偏差の大きさから考察したい。

	ね				い			
名前	声量	ハーモニー	表現の工夫	歌詞の理解	声量	ハーモニー	表現の工夫	歌詞の理解
sum	185	188	169	162	168	168	170	161
avr	4.7436	4.8205	4.3333	4.1538	4.4211	4.4211	4.4737	4.2368
標準偏差	0.8497	0.9140	0.8686	0.8124	0.9482	0.9482	1.0587	1.1255

桃色のセルは標準偏差の小さなもの、黄色、黄緑色は大きなものである。

絶対評価の四項目は六段階評価である。従って、標準偏差が1あるということは平均値の前後1の範囲に68％のデータが存在すると言うことで、例えば平均が4なら3～5の間に七割弱のデータがあり、6や1、2には三割程度しか存在しないと言うことである。六段階評価における標準偏差1というのはかなり大きな散らばりだと考えたい。

ねといは順位を点数に換算したときはね17、い16、とねが上位だった。一方、上記の表を見るように、項目毎の平均を見ると「声量」「ハーモニー」はね、「表現の工夫」「歌詞の理解」はいが上回っており、優位にある項目数では優劣が判定出来ない。

ここで注目したいのが標準偏差で、ねの評定は三項目が標準偏差0.9を下回っている。つまりデータが集中していると言うことになる。一方いは二項目で標準偏差が1を上回っている。そしてその二項目こそがいが優位を示した「表現の工夫」「歌詞の理解」なのである。

絶対評価による評定は、評価者の質を意図して選択したのでない限り、分散の少ないものの方がデータとしての信頼性が高いと考えるべきではあるまいか。もちろん、「この程度の差」がどこまで有意なものであるかはこのデータだけでは判断出来ないが、今回の調査においてはねの評定は信頼度の高いものが揃っており、その上で暫定的ではあっても点数化したときにポイント数で上回ったのである。絶対評価における項目評定を総合した場合、ねが二位、いは三位、という順位が妥当、と考えられる根拠はここに求められる、と考えられる。

もう一点、め、もの項目評定も同じ視座から解釈が可能である。

	め				も			
名前	声量	ハーモニー	表現の工夫	歌詞の理解	声量	ハーモニー	表現の工夫	歌詞の理解
sum	183	147	139	139	157	169	137	133
avr	4.8158	3.8684	3.6579	3.6579	4.1316	4.4474	3.6053	3.5000
標準偏差	0.8336	0.9349	1.0208	0.9939	0.9911	1.2236	0.9737	1.0591

項目毎の優位を見れば、めは三項目で優位となり、点数化された順位ではもよりも上位（め四位、も五位）となったが、一つだけもが上回る「ハーモニー」において、めを大きく上回っている（差は0.5789）。項目数は少なくとも、大きく差がある項目があるならば考慮すべきではないか、と考えることもできる。しかし、ここでも標準偏差に注目すると、もが上回る「ハーモニー」は標準偏差が1.2236とかなり大きい。評価が割れている、ということが見て取れる。本来なら、棄却検定で「とても良い点」「とても厳しい点」が切り捨てられると考えれば、この項目の突出を根拠にもの順位を上方修正することは避けるべきであろう。

と、良く注意すれば実はこの両チームにはもう一つ、大きく評定に差が出ている項目がある。「声量」はもよりもめが0.6842上回っている。さらに、この項目の標準偏差は1未満で両チームとも「妥当な評価」と解釈出来る。つまりは、めが評価項目の数でも三項目と上回り、その突出の数は一項目ずつで同じだが質に於いて信頼度の高い項目で大きな差をつけているわけで、め四位、も五位というのは妥当、と考えられる。

以上が、標準偏差に着目して、絶対評価を基に暫定的に算出した順位を「妥当」であると判断した理由になる。

3-2　絶対評価と相対評価とに生ずる順位の相違、揺れ

2-2で述べた通り、相対評価の平均に基づく順位はこ→ね→い→も→め→やであり、絶対評価をもとに得られた仮順位、こ→ね→い→め→も→やとは四位と五位のめ、もが一致しない。見方を変えれば、それ以外は一致するのだから、順位を出すにあたって絶対評価を用いても相対評価を用いても大きな差はない、とも言えるが、ここでは「差」が如何に生じたか、個々の評価者の傾向から考察したい。

今回の調査では、前稿（第4稿）の相対評価とは大きく手法を変えている。前回は各項目（音程、ハーモニー・バランス、表現の工夫、歌詞の理解）それぞれを「必ず一位から四位まで（第4稿は五グループで調査）」順位付ける、というものであった。今回は「上位であるチームには多くの点を、下位のチームには少ない点を」と点を割り振る方法を採り、なおかつ持ち点を使い切る、というものである。持ち点15は一位から五位までを5、4、3、2、1と配点することを念頭に置いたためである。今回も絶対評価の仮順位を算出するためにこのやり方を使用しているが、一位と二位の差が大きいとき、その評定の差が一点でいいのか、もしくは同順位が複数いた場合の配点をどうするのか、といった問題に対応できるやり方ではない。具体的には、絶対評価「歌詞の理解」の仮順位は五位と六位が同値となり、五位2点、六位1点を平均して両チームとも1.5としているが、最終的に各項目の総和で総合順位を決めるのだから、この部分を場当たり的に処理しては順位の信頼性が低下する。

22344	12
12345	10
23334	7
22335	3
13344	2
22245	2
33333	2
12444	1
13335	1
22227	1
22236	1
11256	1
12246	1
12336	1

相対評価の配点パターンは15種。これを分類すると、

・重複する評定がないもの　12345　一種

・一つの評価が二つ重複するもの　11256、12246、12336　三種

・一つの評価が三つ重複するもの　23334、22245、12444、13335、22236　五種

・一つの評価が四つ重複するもの　22227　一種

・一つの評価が五つ重複するもの　33333　一種

・二つの評価が二つ重複するもの　22344、22335、13344　三種

なお5，10と二チームだけに評点をつけた異常値が一つあったが、計算には含めている（データの棄却を行わないという前提であるため）。

46の評価のうち、重複しなかったのは10データのみ。3/4以上が「同順位」のチームを想定して評価しているのが特徴である。

この15種（異常値を除けば14種）は更に以下のように分類が可能である。

A：上位一チームが突出している→22227(1)、22236(1)、12336(1)、12246(1)、13335(1)、22335（3）　六種8データ

これは最上位と次点とが連続した数値で表されていないものになる。

B：上位二チームが突出している→22245（2)、12256（1)、12246（1）　三種4データ

これは二位と三位が連続していないものになる。12246は同時にAにも分類される。

C：上位三チームが突出している→12444（1）　一種1データ

三位と四位が連続していない（同時に上位三チームは差がない）

D：全体に差は殆どない→33333（1)、22344（12)，23334（7）　三種20データ

このパターンの採点が最も多いデータ数を示した。評定値1と5を両方とも使用しないパターンになる。

なお、異常値もAかつBのパターンを極端にしたものだと言える。

データ数を見ると、Aは22335だけが3データあるが、ほかは1データずつで「少数派」の評価になる。Bも同様に3データある22245以外は少数派。一つのチーム（これは全ての評価においてことなる）だけが突出する、という評価は、相対評価においては少数派である。以下、少数派であるA、B、Cの相対評価と絶対評価の相関を検証してみる。

[22227]

	こ					ね					め					や					い					も					相対					
	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	こ	ね	め	や	い	も
22227	6	6	[illegible]	4	21	0	0	0	0		[illegible]	3	3	3	14	3	4	4	[illegible]	16	[illegible]	3	3	4	15	3	3	4	4	14	[illegible]	0	2	2	2	2

この評価者は、各チームの項目評定の合計で順位を出している。21点のこが他の14,

16，15，14に比べて群を抜いているのは確かであるが、大きな差が生じているのは「ハーモニー」に起因しているのがわかる。また、こ以外の四チームには差があるが、相対評価では差をつけていない。

［22236］

	こ					ね					め					や					い					も					相対					
	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	こ	ね	め	や	い	も
22236	6	6	6	6	24	[illegible]	[illegible]	4	4	18	[illegible]	3	4	4	16	4	3	4	3	14	0	0	0	0		4	[illegible]	3	3	15	6	3	2	2	0	2

この評価者も、22227同様、項目評定の合計に基づいた順位付けになっている。ただし、二位以下の18，16，14，15の18だけを評定3として残りの三チームの評定2と差別化している。一位のこが全項目評定6なので相対評価でも6となっているのだが、「声量」「ハーモニー」の両項目が揃って高評価とならない場合、相対評価に反映しないという特徴がある。

［12336］

	こ					ね					め					や					い					も					相対					
	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	こ	ね	め	や	い	も
12336	4	3	4	4	15	3	3	4	3	13	4	2	2	2	10	0	0	0	0		4	4	2	3	13	3	4	2	2	11	6	3	2	0	3	1

この評価者は絶対評価において高い評定値（5や6）を一つも使っておらず、厳しい評価者だと言える。相対評価はそのまま項目評定の合計の高い順になっているが、15のこと13のね、いとがこれだけ大きな差になる（こ→6、ね、い→3）根拠はこの採点結果からは伺えない。

［12246］

	こ					ね					め					や					い					も					相対					
	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	こ	ね	め	や	い	も
12246	0	0	0	0		6	[illegible]	[illegible]	[illegible]	21	4	2	3	4	13	4	2	3	3	12	3	[illegible]	[illegible]	[illegible]	18	1	2	1	1	5	0	6	2	2	4	1

この評価者は下位チームに厳しい採点となっている。21，13，12，18，5であるから非常に分散の大きな採点であるが、相対評価は上位に大きく点を出し、下位は大きく点を下げてはいない。持ち点制度で相対評価を行うと、上位に意を尽くす反面、下位は細かな成果を無視される、という例だと言える。仮に持ち点が20くらいあれば、三位、四位に差をつけた上で五位をもっと低く評価した可能性がある。

［13335］

	こ					ね					め					や					い					も					相対					
	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	こ	ね	め	や	い	も
13335	4	4	4	4	16	4	4	3	4	15	4	4	4	3	15	3	2	3	3	11	4	4	3	3	14	0	0	0	0		3	[illegible]	3	1	3	0

中間順位の三チームを同点とした評価だが、絶対評価を一切考慮せずに相対的な順位を出している貴重なサンプル。項目評定の合計では16のこが一位だが、15のねが一位と評価され、同じ15のめ、14のい、16のこが同着二位となっている。何がこの相対順位の根拠となったかは、絶対評価の項目評定から読み取ることができない。

[22335]

	こ					ね					め					や					い					も					相対					
	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	こ	ね	め	や	い	も
22335	6	[illegible]	3	3	17	0	0	0	0		4	3	2	3	12	3	2	2	3	10	3	4	4	2	13	4	[illegible]	3	3	15	[illegible]	0	2	2	3	3
22335	6	[illegible]	[illegible]	3	19	4	4	3	3	14	[illegible]	4	4	3	16	0	0	0	0		[illegible]	[illegible]	4	3	17	4	4	4	3	15	[illegible]	3	3	0	2	2
22335	6	6	6	[illegible]	23	4	[illegible]	4	4	17	[illegible]	[illegible]	4	4	18	4	3	3	3	13	0	0	0	0		4	3	3	3	13	[illegible]	3	3	2	0	2

三名ともこを一位としており、最下行の一人を除くと、二位三位、四位五位が同点となる理由は項目評定の合計では読み取れない。

[22245]

	こ					ね					め					や					い					も					相対					
	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	こ	ね	め	や	い	も
22245	0	0	0	0		4	3	4	4	15	3	3	2	2	10	4	3	3	3	13	4	4	3	3	14	4	3	2	2	11	0	[illegible]	2	2	4	2
22245	6	6	[illegible]	4	21	[illegible]	[illegible]	[illegible]	4	19	4	3	3	3	13	0	0	0	0		[illegible]	4	4	4	17	6	[illegible]	4	3	18	[illegible]	2	4	0	2	2

上位二チームが抜けており、下位三チームは差がない、という評価。抜きん出た二チームは項目評定の合計から導かれたと思われる一方、下位三チームが「同じ評定」であることの理由は不明。

[12256]

	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	こ	ね	め	や	い	も
11256	6	4	[illegible]	6	21	[illegible]	6	6	[illegible]	22	0	0	0	0		4	2	2	3	13	4	3	4	2	13	[illegible]	3	4	3	15	6	[illegible]	0	1	2	1

この採点は非常に興味深い。基本的に項目評定の合計に従っているが、一位と二位は合計には従っていない。21のこが一位、22のねが二位。

[12444]

	こ					ね					め					や					い					も					相対					
	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	声量	ハーモニー	表現の工夫	歌詞の理解	計	こ	ね	め	や	い	も
12444	6	6	[illegible]	6	23	4	6	4	4	18	0	0	0	0		2	2	4	4	12	6	[illegible]	6	4	21	[illegible]	6	4	3	18	4	2	0	1	4	4

上位三チームが同評価、というものだが、項目評定の合計を見る限りこ、い、もの順に順位がつくはずなのに、同評価。さらに、もとねは18ポイントで同点だが、ねの評定は五位相当の2。つまり絶対評価が全く相対評価に反映していない。

以上、持ち点制度を活用することで個性的な採点をしたものを個別に検証すると、「相対評価による順位は絶対評価による項目評定とは必ずしも相関しない」ことが見えてくる。

次に、持ち点制度であるにも関わらず、順位点と同じ12345という評定を付した10人を検証してみる。

[12345]

こ					ね					め					や					い					も										
声量	ハーモニー	表現の工夫	歌詞の理解		声量	ハーモニー	表現の工夫	歌詞の理解		声量	ハーモニー	表現の工夫	歌詞の理解		声量	ハーモニー	表現の工夫	歌詞の理解		声量	ハーモニー	表現の工夫	歌詞の理解		声量	ハーモニー	表現の工夫	歌詞の理解		こ	ね	め	や	い	も
6	4	[illegible]	[illegible]	20	[illegible]	6	[illegible]	[illegible]	21	6	3	3	3	15	4	3	3	3	13	0	0	0	0		4	[illegible]	4	4	17	[illegible]	3	2	1	0	4
6	6	6	6	24	6	6	[illegible]	[illegible]	22	6	4	[illegible]	[illegible]	20	4	3	4	4	15	0	0	0	0		[illegible]	6	[illegible]	[illegible]	21	[illegible]	4	2	1	0	3
6	6	6	6	24	0	0	0	0		6	4	4	4	18	[illegible]	4	[illegible]	[illegible]	20	[illegible]	[illegible]	[illegible]	[illegible]	20	4	[illegible]	[illegible]	[illegible]	19	[illegible]	0	1	3	4	2
6	6	6	6	24	6	[illegible]	0	[illegible]	21	4	4	3	4	15	0	0	0	0		[illegible]	6	[illegible]	6	22	[illegible]	6	3	4	18	[illegible]	4	1	0	3	2
6	6	6	6	24	4	4	3	3	14	[illegible]	4	[illegible]	3	17	3	2	2	2	9	0	0	0	0		3	3	2	2	10	[illegible]	3	4	1	0	2
6	[illegible]	[illegible]	4	20	[illegible]	4	3	3	15	0	0	0	0		3	1	2	2	8	3	3	3	3	12	3	2	3	3	11	[illegible]	4	0	2	3	1

こを一位とした六名は一位のみ絶対評価の項目評定合計に沿っているが、二位以下は一行目がね21（三位）も17（二位）、三行目がい20（三位）や20（四位）、六行目がや8（五位）も11（六位）と、六例中二例で順位逆転が、一例で同点の無視が起きている。

こ					ね					め					や					い					も											
声量	ハーモニー	表現のエチュード	歌詞の母音		声量	ハーモニー	表現のエチュード	歌詞の母音		声量	ハーモニー	表現のエチュード	歌詞の母音		声量	ハーモニー	表現のエチュード	歌詞の母音		声量	ハーモニー	表現のエチュード	歌詞の母音		声量	ハーモニー	表現のエチュード	歌詞の母音		こ	ね	め	や	い	も
6	3	0	5	19	0	0	0	0		4	6	1	1	12	5	2	3	3	13	5	4	6	6	21	5	2	2	2	11	4	0	2	3	5	1
6	4	5	6	21	0	0	0	0		5	3	4	4	16	5	3	3	5	16	6	6	6	6	24	6	6	5	5	22	3	0	1	2	5	4

いを一位とした二名ではなぜか順位逆転は起きていない。二行目同点のめとやに同点の無視が起こり、差がつけられている。

こ					ね					め					や					い					も											
声量	ハーモニー	表現のエチュード	歌詞の母音		声量	ハーモニー	表現のエチュード	歌詞の母音		声量	ハーモニー	表現のエチュード	歌詞の母音		声量	ハーモニー	表現のエチュード	歌詞の母音		声量	ハーモニー	表現のエチュード	歌詞の母音		声量	ハーモニー	表現のエチュード	歌詞の母音		こ	ね	め	や	い	も
4	4	5	4	17	0	0	0	0		3	3	3	3	12	3	2	2	3	10	3	3	4	3	13	4	5	4	5	18	4	0	2	1	3	5
0	0	0	0		3	5	4	4	16	4	2	3	3	12	3	2	2	2	9	3	4	5	4	16	3	4	3	3	13	0	4	1	2	3	5

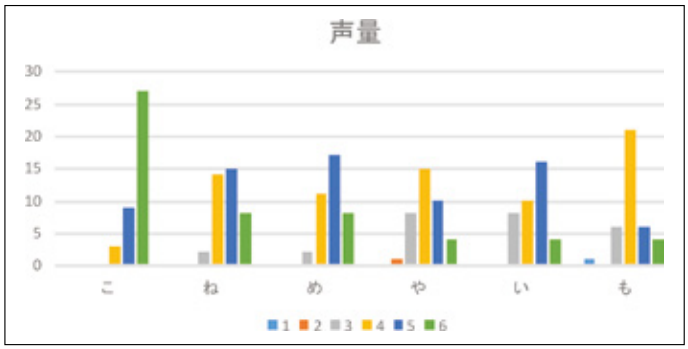
もを一位とした二名は二行目で大胆な順位逆転が起きている。ね16（二位）い16（三位）も13（一位）。こういうことは相対評価でしか起こらないと結論できる。

おそらく、[12345] という評定の割り振りを採用した評価者は、この評定の割り振り有りきで採点したのだと推測する。だからこそ、項目評定の合計が同点でも順位を変えているのだと思われる。ただし、大胆な「逆転現象」が何故起こっているのかは、このデータからは推測することができなかった。

3-3　絶対評価の妥当性

相対評価による順位付けに必ずしも「絶対評価に基づく項目評定」が関連していないのであれば、絶対評価の項目評定は音楽表現の優劣を判断する際に用いないほうがいいのであろうか。

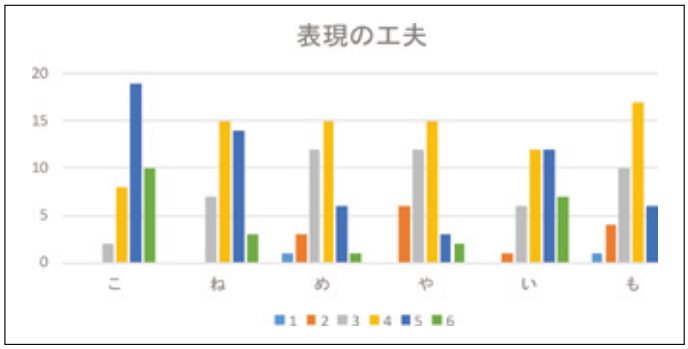
以下に示すのは、各チーム毎の絶対評価の各項目評定分布を示した棒グラフである。項目ごとに順位を考えるとき、評定値の合計を用いることに疑問があるとしても、この分布を見る限りチームごとの特徴は如実に現れていると結論できる。



「声量」で注目すべきは最頻値（モード）であり、各チームごと最もグラフの高いものを見るとこが6、ね、め、いが5、や、もが4となる。ね、め、いの三チーム順位を見るなら二番目、三番目に数値の高いものを同時に注目すれば、ね、めはほぼ差がなくいは若干ね、めに及ばないことがわかる。や、もは評価の一致するも、

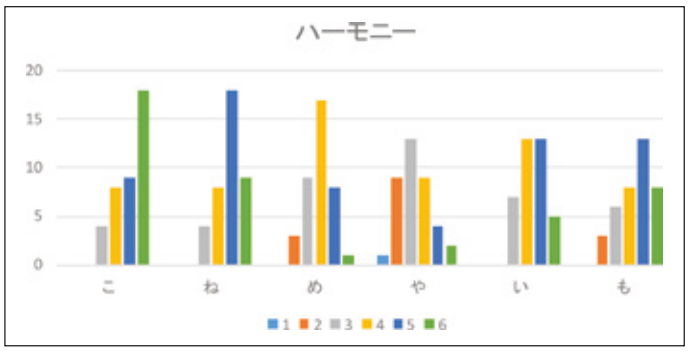
評価の分かれるや、と捉えるのが妥当であろう。つまり、絶対評価の項目評定は、「視覚化」することでイメージとして把握する利用法が推奨されるべきだと考えたい。

同様の傾向は他の項目からもうかがえる。「表現の工夫」は以下のようになる。

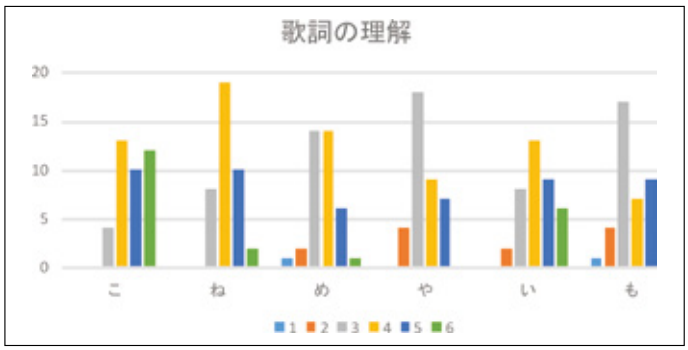


6と5に集中するこ、6がこに次いで目立ち5，4も高いい、5，4はいよりも高いが6で劣るね、この三チームには順位差が見える一方、め、や、もには明確な差異が見えない。この「“差異が見えない”ことが見える」ことが重要だと考える。つまり、絶対評価は数値化する必要はない、と考えたい。

これは「ハーモニー」でも顕著に読み取れる。この項目はモードを見るだけで6が多いこ、5が多いね、4が多いめ、3が多いや、この四チームの出来栄えの差はほぼ明瞭である。ただし、い、もの評価は簡単ではない。モードを見ればいは4，5でもの5と同じレベルにあるが、もは6がほぼねと同じレベルで多い。更にもは2の評定が多く、六チームの中では最も評価がばらついている。こうした場合は、いやもの優劣は数字で示すべきではなく、あくまで図示されたものから読み取り、個々に解釈すべきだと考えたい。



「歌詞の理解」は更に数値化することの無意味が見えるデータとなっている。モードだけに注目すれば4がこ、ね、め、い（めは3と同値）、3がめ、や、も（めは4と同値）にピークがあるが、いわゆる正規分布になっているね、や、い、めと異なり、こ、もはダブルピークの分布になっている。この項目はこうした分布を視覚化すること以外に評価の意味を伝える方法はないと考える。6の数を見ればこ、いが多い



一方、め、い、もは評価がばらついている。評価がばらついていないこ、ね、やではこがダブルピークを示している。も、やはどちらが優れているのかこのグラフから解釈するのは難しい。

4. まとめ

本研究は前稿で得られた以下三点の結論を基にしている。

1　相対評価は順位決定に、絶対評価はフォローアップに効能あり
2　データのばらつきの大きさは、絶対評価では起こりがちだが、相対評価で起こることは稀であり、両者が連動している時は表彰の対象とできる可能性あり
3　絶対評価と相対評価の標準偏差がともに大きい場合、問題は審査される側にあるとは限らない

1に関してはまったく同じ結論となった。特に絶対評価での順位算出は牽強付会をもたらし、その信頼性も高くない。飽くまで、教育的効果を見据えたものであるべき、と結論する。そして、その与え方は数字ではなく、図、グラフになる事が望ましい。

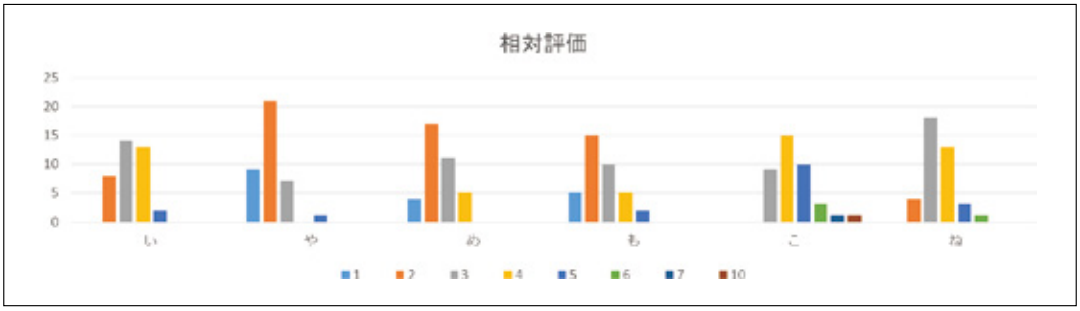
3-3では各項目毎にそれぞれのチームがどんな評定を獲得したかをグラフ化したが、各チーム毎にそれぞれの評定値がどんな項目に多いかをグラフ化するやり方も効果があると考える。このようにまとめると（2-1参照)、チーム毎の演奏の質がかなりはっきりと見えるので、特に教育的意味は大きい。

2の「評価のばらつき」は標準偏差拡大という目標を持って絶対評価、相対評価ともいくつかの工夫を試みたが、項目の名称を「音程」→「声量」と改めたものについては、評価項目としてはあまり意味がなかったと考える。とはいえ、実際に学校現場で項目別評価をするとき最も明確に判定されるのもこの「声量」である事は間違いないのだから、この試み自体に意味がなかったとは思わない。今回は46名の評価者が（前回の研究と同様）かなり“的確”な判定をしていたため、特に「声量」項目では評価がそれほど散らばらなかった。これには今回、録音の環境が改善したことも大きかったと考える。

3の標準偏差の相関、に関しては、今回の相対評価が持ち点制度であったことと、参加六チームが非常に質の高いボイスアンサンブルを完成させたことによって採点に大きな差をつけたものが少数派となってしまい、標準偏差の拡大という目標は達成できなかった。その代わり、差の大きな採点をしたもの、非常に保守的な採点をしたものの採点傾向を絶対評価・相対評価で比較したことによって、絶対評価の項目評定を合計した数値は順位算出には用いるべきではないと結論できた。

今回の調査で得られた新たな結論は、相対評価による順位を如何にして「使うか」である。

一人の評価者が全てを決めるのならともかく、複数の評価者の出した順位をどう総合的に扱うべきか。



これに関しては“最頻値利用”を提言したい。2-2のグラフを再掲する。

上のグラフは、チーム毎に相対評価のポイント毎に何人の票が入ったかを示したものである。グラフ上の高さが最大になるのは2→や、め、も、3→い、ね、4→こ、であり、これだけで三つのカテゴリに分かれることは自明である。2が最頻の三チームでは2より大きな評価の獲得状況を見ればも→め→やであることはすぐに納得できる。3が最頻のい、ねは、このグラフからわかることは「どちらが上でも僅差」ということだろう。つまり、最頻値に注目する、というグラフの読み方をしっかりと指導して、数値ではなく図表で提示…という絶対評価と同じ手法がこちらでも有効と結論する。

特に推奨したいのが「グループ分け」で、上中下とか松竹梅のような判定群を設定し、そのどこに入るかくらいで留める方が相対評価による順位算定には相応しいのではあるまいか。同じ群の中での優劣はある意味鬩ぎ合うことを期待出来る一方、群を隔てる壁は高い、と受容を求める方が建設的であると感じる。

この群は三群くらいに分けるのが最も適切だと考えるが、場合によって二群や四群にせざるを得ないことも生じるだろう。二群制だと一軍二軍、主力と補欠、のような主と従になってしまう可能性が懸念されるが、群を隔てる壁を明確化出来る利点がある。四軍以上になれば今度は優れる、劣るではなく、特質の差としてグルーピングするのも一つのやり方になる。たとえば相対評価は三群制で発表し、絶対評価は項目別四群制で発表する、などの手法が考えられよう。

論を閉じるにあたって、新たに見えてきた課題について触れる。

今回は、模擬合唱コンクールを課題曲、自由曲の二曲で競ったが、評価は二曲を総合したものだった。特に、絶対評価を課題曲と自由曲で分けると、音楽表現上の優れた部分や、問題が更にクリアになったのでは、と感じた。課題曲は全チームが演奏するので、自チーム以外の課題曲の演奏を評価するとき、「演奏を経験した曲」として評定が出せることを評価の分散増大に生かせたのではないか、と考える。加えて、“自チーム以外の五チーム”という縛りを撤廃して、自チームも採点させれば自己評価と他者の評価とを比較すること

で評価者としての信頼度を確認できたのではあるまいか。こう考えるのも、総じて評価者の出した評点が的確だ、と感じたためである。もう少し「好み」に依った極端な評価が出ることを予想していたが、どう考えても異常、という評点は相対評価に一つあっただけ(それもルールを勘違いした方向のもの)であった。この「異常値」が顕在化しなかったことも、データの棄却に関する考察が進まなかった一因となる。そもそも評価者の数が50に迫るからこそ「データ棄却」を検討するわけであり、数人の評価者が順位を決める学校でのコンクールには「上下カット」のような棄却はなじまないのかもしれない。

相対評価には新たな「持ち点制」という手法を導入したが、こちらも課題曲と自由曲を分けた上で、総合を出すことを求められたのではないだろうか。特に課題曲は評価者が多くなればその評価の妥当性の検証は確実性を増すので、本研究のごとく「参加者がそのまま評価者になる」という、多くの評価者の出す評定を総合化するならば一考の余地があった。

勿論、それを行わなかったことには理由があり、一言で言えば「全参加者に採点してもらう」ために、採点することから煩雑さをなるべく排除しようとしたからである。特に後日配信される動画を視聴しての採点なので、時間は演奏時間分は必ずかかる。リアルタイムの演奏を聴きながら、ワークシートのような採点用紙に書き込むならば、もう少し項目を多くすることは出来るのかもしれない。

絶対評価の分散を大きくすることには成功していない。第4稿で課題としたことが再び課題になってしまった。六段階より段階を増やし、八段階、十段階にすればもう少し分散は大きくなると予想できるが、逆に低い評価は出にくくなる。絶対評価なので持ち点制は不適であり、また、評定毎に個数制限を設けるのも確かに分散は大きくなるが絶対評価の利点を削いでしまう。課題として次回以降に委ねたい。

次回以降に申し送るかどうかとは別に、課題として現在考えているのは「評価者数が少ないときに、絶対評価・相対評価をどう援用すれば音楽演奏の評価に役立つか」になる。回数を重ねることで一連の「評価に関する研究」は、評価者の質が揃っていることが利点であると考える。ただし、実際の音楽演奏評価…特に学校主催の合唱コンクール…ではそれは必ずしも期待できない。「持ち点傾斜制」はそうした評価者の質が保証されない時の対応策だと考えているが、結果的に音楽演奏以外の要素を評価して順位が出ていることも否定は出来ないだろう。教育者としての評価眼があれば判定できる要素であるならば、その評価の質は保証できことになるので、理解できる方法ではある。しかし、個人的には「音楽的要素だけを」多くの評価者で的確に判定することを理想としたいと考えている。

評価項目としてどんな項目立てをするかは一種の諸刃の剣である。たとえば、声量は非常に信頼度が高い項目であることが明らかになったが、これは逆に危険で、誰にでもわかる判定要素が異常に強調される可能性がある。とはいえ、実際の合唱において「声が大き

いか小さいか」は一種の必須項目になっており、これをクリアしなければこれ以降の審査にかからないことも否定できない。

この関連になるが、バランスの項目を設定すべきだった、というのは反省点である。ハーモニーとして評価してくれることを期待したが、いくつかのコメントを見る限り「バランスという項目が欲しかった」との要望は強かった。ハーモニーやバランスは音楽的力量の高さを要求する項目なので、評価者の数を増やした時、全評価者に要求するのは本来困難なのだが、本研究においてはむしろなくてはならない項目であった。

チーム成員の人数を少なくした反動で声楽専攻のアドバンテージが強調されてしまったことは否めない…圧倒的に一位として推されたこであるが、七名中五名が声楽専攻で声が大きく美しいのは当たり前であった。ただし、技量や経験に優れたチームがある事自体は非難されるべきことではなく、そうした優位性を持たないチームの良さを如何に評定値に落とし込めるかが重要だと考える。上手な演奏を上手であると判定できたことは寧ろ成功だと考えている。

やはり、問題は「順位」、どのチームに較べてどのチームが優れるのか、である。AがBに優れると言うことは、BがAに劣ると言うことも同時に意味する。順位は、価値が多いか少ないか（多寡）、を示すものではなく、価値の質や好みを示すものであり、これを音楽演奏を競うことによって理解する。学校で行われるコンクールとはここを目指すべきだ。

絶対評価の項目評定において演奏の多面的価値を示し、審査員という「音楽力に優れた個人」の理解において価値づけられたものとして相対的順位を示す。理想がここにある事から揺らぐことなく、今回解決できなかった諸問題を次稿で継続して取り上げたい。

現在「やったら面白そうだ」と感じるのは音源による評価の揺れである。今回、音響スタジオで録音したことにより「高音質の録音データ」が得られたが、録音環境によって様々な要素が底上げされたことに鑑みて高音質データは審査に用いていない。現場音響とスタジオ音響、二つの音源の差を聞き分けられることと採点の信頼性との相関関係が調査出来れば、リモート採点などの時と場所を共有しない採点方法への後押しになる可能性がある。新たな視座として検討したい。