



## National Twelve-Bell Striking Contest

### Defining Good Ringing (or to be more precise, how to define imperfections in ringing)

#### Survey Results

Thank you very much to all 29 participants for your thoughtful and detailed responses.

Participants:

Jonathan Agg	Phil Barnes	Matthew Beadman	Mark Bell	Graham Bloom
Wendy Bloom	Tim Bradley	Alex Byrne	Martin Cansdale	Julia Cater
David Dearnley	Jennie Earis	Mark Gill	Tom Griffiths	Tom Hinks
John Hughes-D'Aeth	Chris Kippin	Colin Lee	Simon Linford	Ed Mack
Tom Mack + other SRCY	Paul Mounsey	Chris Poole	Ian Roulstone	Philip Saddleton
Katie Town	George Unsworth	Stef Warboys	Lucy Woodward	

So that you can follow an individual's answers, each respondent has been assigned a number, R1, R2, etc. The numbers have been randomly assigned to preserve anonymity.

To preserve anonymity, we have blanked out a few words where for example a respondent has written "when I judged in xxxx" or "when we won in xxxx". This only affects 3 or 4 sentences in the entire document.

Thank you very much from the Working Group (Neil Buswell, David Pipe, Chris Poole, and Stef Warboys).

1. What are the key metrics you'd look for in defining good ringing, and how important are they? Please mark each with a score of 1-5.

	1 (not very important)	2 (less important)	3 (doesn't massively matter)	4 (important)	5 (very important)
<i>Average speed for the touch</i>	*****	***	*****	**	
<p><b>4-star replies:</b>            R13. While judging should to some extent be speed agnostic palpably slow ringing on twelve never sounds pleasant even when accurate.            R14. Important only in as far as it produces a nice sounding piece of ringing - i.e. not sounding rushed or stilted, rather than conforming to an ideal speed.</p> <p><b>3-star replies:</b>            R12. Needs to be within sensible bounds, which depend somehow on the "sound" of the bells.            R28. Even though ringing speed should not be judged, the speed of the ringing may be a contributing factor in whether the ringing was good or bad and this would be more of a general comment about the ringing from the judges rather than a cause of faults.</p> <p><b>2-star replies:</b>            R16. As long as the striking is good, speed is not particularly important, though some bells may have a more optimum speed at which it is easier to achieve good striking            R19. Within reason – for a given ring there is normally a wide range of acceptable speeds.</p> <p><b>1-star replies:</b>            R15. Accuracy trumps speed            R20. As long as a band isn't in at extremes speed, I don't think is relevant.            R24. There may be speeds for the touch that are not appropriate for the band. I remember ringing in an eliminator, I think, at xxxx. We rang very well and very accurately in a very controlled way. The local band's verbal response was "you can't ring those bells at that speed". Our response was "we just did" – and it was an emphatic victory.            R27. Speed will certainly influence the quality of the ringing (as a function of the interaction of bells, band etc), but I do not believe we should be defining the quality of ringing itself by the speed. I.e. in particular in comparing 2 pieces of ringing, whilst we may well hypothesise that the quality of one particular piece may be affected because the 'chosen' speed is causing problems, it should be the result that is considered/important in assessing the actual quality, not the reasoning.</p>					

<b>Consistency of speed across the touch</b>	**	*	*****	*****	
<p><b>4-star replies:</b></p> <p>R12. OK for this to change gradually, provided everyone does it sufficiently together and it feels seamless</p> <p>R19. A few of these metrics are not necessarily things that you would score bands on, but are contributors to the overall standard of striking – if the speed changes but the striking remains accurate that is not a problem.</p> <p>R25. A bit more important because unlikely that all bells will respond to speed changes equally well.</p> <p><b>3-star replies:</b></p> <p>R16. Ditto, though will be even more important to helping to achieve good striking</p> <p>R27. I am undecided on this point. Overall I think probably similar comments to speed. Whereas different bands may (legitimately) be aiming for a particular (/different) speed, I would imagine that most people would suggest it is desirable to aim for consistency of speed (albeit again as a method of achieving/assisting quality ringing, rather than the consistency itself being a defining factor). Indeed it may be that the difficult-to-define ‘artistic’/‘easy-on-the-ear’ factor <i>could</i> be improved by natural variations in speed if the whole band goes with any speed changes consistently. That said, compounding this would be the rate of change: inconsistent speed in terms of individual rows rather than more gradual speed changes across the touch may be more important (although again I would hypothesise would <i>result</i> in more inaccurate individual blows (cause and effect), so there should be no need to consider this directly in terms of <i>defining</i> quality).</p> <p>R28. As above.</p> <p><b>2-star replies:</b></p> <p>R20. In itself I don’t think this is relevant. A preference for dynamism or consistency is subjective, and valid either way, though of course inconsistency is often part of the cause of poorer ringing</p> <p><b>1-star replies:</b></p> <p>R15. A consistent rallentando (for instance) across the touch by the entire band is a specific skill. If done well it should be applauded, not punished. This is a performance, after all.</p> <p>R23. Although speed can affect how hard it is to achieve consistency, I don’t think good ringing is dependent purely on how fast or slow a band rings.</p>					

<b>Consistency of leading (with a handstroke gap), denoted by a standard deviation</b>				*****	*****
<p><b>5-star replies:</b></p> <p>R22. What is meant by a handstroke gap? The key thing is that every bell does the same thing</p> <p>R23. In our team we talk about rhythm and separation occurring inside the boundaries of a ‘box’ which I find helpful, and I think emphasises the centrality of consistency in leading and 12<sup>th</sup> position, with consistency of gaps between the other bells within this box.</p> <p>R24. Denoted by a standard deviation is not relevant – it just needs to be one beat of 25. These comments apply to the subsequent questions too. I have graded my response by how important this is. Note by means of observation and experience of judging lesser competitions: Ability to lead accurately is probably the key differentiator between ‘a good’ ringer and ‘an excellent’ ringer</p> <p>R28. Consistency is a key factor in good ringing.</p> <p><b>4-star replies:</b></p> <p>R12. How the previous backstroke was rung can make the same sized gap feel and sound very different</p> <p>R20. Unevenness will in my view be quite unpleasant to listen to; audience and ringers alike enjoy feeling a rhythm, which inconsistency of leading disrupts.</p> <p>R27. Clearly important in defining the change, setting the tone and also general striking, and will influence the ability of the rest of the band to achieve good striking. I am not wholly convinced that in terms of defining the <i>quality</i> of a piece of ringing that this should be considered any more important than any other individual ‘blow’ in the row. (Cause or effect?). In contrast to tenor ringing (below) – in terms of actual/perceived quality of the sound/striking (in contrast to <i>enabling</i> good striking across the band) I think the leading/position of the leading does more specifically define the change than any individual bell (albeit there will be a strong interaction here with consistency in length of change as above – though again, whilst it will potentially be important to an individual band in considering how to <i>improve</i> the overall result, I am not convinced that in itself it should be heightened in importance over any other blow in defining the actual quality)</p>					
<b>* Consistency of tenor ringing, denoted by a standard deviation</b>	*	*	*	*****	***
<p><b>5-star replies:</b></p> <p>R13. Commenting on the asterisked footnote. Is there proper statistical evidence that Strikeometer analysis from HawkEar data does treat all bells equally – specifically the tenor?</p>					

**4-star replies:**

R12. Should be setting the rhythm, but can be made up for. Much easier to pick out for those of us without perfect pitch than other bells

R14. No more important in striking terms than other bells, but inaccuracies here will make good ringing more difficult.

R20. If the tenor is striking unpredictably this will be unpleasant. But if the band is moving with the tenor this will be fine.

R23. Important, but not an independent variable within a regression – only a predictor of good ringing when consistency occurs in the context of other bells ringing consistently.

R24. Denoted by standard deviation – why is this in all the questions in this section? It should be measured by where ‘the right’ place is – on a 25-beat rhythm. As I understand it, SD does not measure how ‘right’ it is, only ‘how consistent’ it is. These comments apply to the subsequent questions too. I have graded my response by how important this is. Excellent tenor ringing is very important. It is more important than excellent ringing of, say – for example, 3 or 4 because it makes the loudest noise and is the easiest bell to ‘tune into’. It is also (for the same reasons) the easiest bell to be distracted by...and maximises the number of team members likely to be ‘thrown off’ by bad blows, mistakes or consistent error. There will be individuals who are more adversely impacted by this type of error being made by other bells but I am generalising on behalf of those who do not possess ‘fantastic ears’. Stedman is a special case here – I’ve limited my comments to maximus but the tenor is even more obvious for cinques – of course. You may get **misconceptions** about how the tenor impacts ‘speed’ in responses to this question.

R25. I’d say perceived as important throughout, but particularly when with other back bells at front or end of change.

R27. Having a consistent tenor will naturally be very important to any band in achieving good quality ringing (in making it easier for the rest of the band to establish a rhythm and clear aim), however, similarly I think it would be a mistake to single out the consistency of the tenor (or any other individual bell) as being more important *per se* in judging the quality of the sound. (i.e. each bell/position should be considered equally in judging the accuracy of the *sound*, regardless of any perceived importance, or otherwise, in enabling the band to achieve that sound).

R28. As above.

**1-star replies:**

R15. Either give consistency scores for all bells, or none. Singling out specific bells makes no sense to me.

**\* Consistency of treble ringing, denoted by a standard deviation**

\*

\*

\*

\*\*\*\*\*

\*\*\*

**4-star replies:**

R12. More important in maximus than Stedman

R14. No more important in striking terms than other bells, but inaccuracies here will make good ringing more difficult.

R20. This could be unpleasant if treble is audibly on occasion hitting bells.

R23. Important, but not an independent variable within a regression – only a predictor of good ringing when consistency occurs in the context of other bells ringing consistently.

R24. Exactly the same comments as for the tenor above, except it's not the loudest – but it's probably the next easiest to hear – although some would argue 11 or 11 /9, when ringing maximus.

R25. I'd say perceived as important at lead ends.

R27. Ditto above.

R28. As above.

**1-star replies:**

R15. Either give consistency scores for all bells, or none. Singling out specific bells makes no sense to me.

<b>* Consistency of not the treble or tenor (i.e. another bell), denoted by a standard deviation</b>	*	**	***	*****	***
--	---	----	-----	-------	-----

**5-star replies:**

R12. Need enough rhythm-setters in the band to build and preserve the framework

**4-star replies:**

R24. See previous comments – the answer is still 4 (but less than the tenor and treble). When I am judging and I have identified a persistent error, say 10 is quick at back, I find it very difficult not to very heavily penalise the error. My brain can't 'stop' hearing the error unless it is almost completely eradicated. I don't know the exact measure of such an error. People tend to say you can only hear an error at approx. 30ms. Our rule of thumb, degree of certainty is that almost all our band can hear (and deal with) 20ms error. I'd love to know how much of a persistent error would need to be eradicated before I started not to 'automatically' hear it. Now, I know that other judges believe they 'stop' hearing or penalising a persistent error unless it is extreme. I believe this is an example of the human brain doing diametrically opposite things to each other and this is a good example of where highly competent human judges and the machine would come out with different answers – and the machine would probably be right!

R27. Ditto above.

R28. As above.

**3-star replies:**

R20. Does matter but not quite as much as the tenor/ trebles.

R25. It's not that it doesn't matter, but perhaps bells other than 1 or 12 are less noticeable, particularly at lead ends and half leads for 1 and roll-up points for 9,10,11,12.

**1-star replies:**

R15. Either give consistency scores for all bells, or none. Singling out specific bells makes no sense to me.

**Consistency of length of change (stringing bells out at the back of the change, or failing to fully reach the back of the change)**

\*

\*

\*\*\*\*\*

\*\*\*\*\*

**5-star replies:**

R22. This question does not need to appear in a survey about good ringing. Neither example is going to produce good ringing.

R23. I think one of the most important metrics.

**4-star replies:**

R17. I enjoy some minute variation where musical changes or rollups are involved

R28. As above.

**3-star replies:**

R27. As a general concept, I think we have come to accept that this consistency is important (vs any potential for artistic *improvement* that may be conceivable possible by e.g. consistent [and even] rhythmical deviation across the change.) This will also clearly be another "cause and effect" factor. I'm not sure that any *increased* importance should be given to this in defining the overall quality of the ringing, over and above the simple fact that any inconsistency would naturally lead to measured errors in individual strike position.

**2-star replies:**

R20. I caution against penalising this if it isn't resulting in clashes or rhythm disruption. Some variability can give a pleasant and more exciting dynamism.

**Non-marked reply:**

R15. I don't think I've grasped what you're trying to get at here. If you mean that having exactly the same overall length for every change is the ideal, I disagree (as noted above). If you mean that having isolated and random bells that 'string out' at the back for instance, I agree.

R24. This drives me mad when I am judging.

**Consistency in  
maintaining the same  
inter-bell gap  
throughout the change**

\*

\*\*\*

\*\*\*\*\*

\*\*\*\*\*

**5-star replies:**

R12. Critical within a change, and needs to be consistent from row to row, but OK to change gradually through the touch (as this changing is clearly the mechanism for the speed changing)

R22. This is fundamental to what is meant by good striking

R23. I think one of the most important metrics.

R24. This is exactly what good striking is. I know there are those who express a preference for roll-ups or certain key changes to be stretched out (savoured) or powered-through (sped up). I expect you can probably find as many ringers who 'prefer' one as the other'. I am as happy with any random change coming up 'perfectly'. You can ring wonderfully well despite being tone deaf and there are some examples of this. Personally, I think people tend to comment on those changes, simply because more people 'recognise those changes'. Ringers with wonderful ears comment on different parts of the test piece. I think judges actually comment on this a bit too much. I would choose to comment about the excellent ringing "between rows x and y, which is as the bells approached the middle".

**4-star replies:**

R19. Note that there are issues with using HawkEar to create metrics based on the inter-bell gap – if this is used to mark faults the quantization means that at certain speeds there is a step-change in the size of the error that will be counted as a fault.

R28. As above.

**3-star replies:**

R18. This links to pace really.



R20. Within in change, and particularly within groups of bells (e.g. a roll up) this is very important. But it's not so important if the inter-bell gaps vary across changes (if executed without clashes).

R27. Ditto above (is this the above not just a subset of this?)

**Non-marked reply:**

R15. Afraid I don't know what an inter-bell gap is.

<b>Counts of overall major faults (noticeable large clips, clashes / noticeable large gaps)</b>			*	**	*****
---	--	--	---	----	-------

**5-star replies:**

R12. These are very disruptive – good ringing can't have many of these

R13. Anything that perturbs the rhythm spoils good ringing. Both major and minor.

R20. This is the most audible way in which ringing can be unpleasant to hear.

**4-star replies:**

R27. Larger individual errors are clearly important to the overall sound (notwithstanding effect on the bands rhythm etc). I am comfortable with the concept that major faults could be considered more important than purely their individual effect on standard deviation/error. However, both the magnitude of any additional importance, and any thresholds for what constitutes 'major' (vs minor etc) will be very subjective (and I would be in favour of this NOT being defined). I appreciate there is an argument that we need to define everything so that everyone has a level playing field, but I also consider in this case, the aim (absence of major faults) is clear, regardless of the perceived importance of individual errors. This is an area where the individual preference of the judge/judging team can and should be allowed for from year to year. (This may/should be possible already, but e.g. could/should hawkear output be set to allow judges to e.g. provide counts for major errors using any input threshold level, rather than just the [current?] 50ms [I think] level?)

<b>Counts of overall minor faults (small, noticeable errors)</b>	*			*****	*****
--	---	--	--	-------	-------

**5-star replies:**

R13. Anything that perturbs the rhythm spoils good ringing. Both major and minor.

R14. As important as larger errors, but carrying less weight per error.

R24. One could get into an argument about what a noticeable error is. When I judge, I only mark a 'perfect' change as 'clean' (i.e. no faults) and I assure you that there are very, very few of those in any of the National 12-Bell winning test pieces. The winning team is the nearest to 'perfect' with major error marked more severely than minor error. Nevertheless, major error almost always causes more disruption to the ringing than minor – because other bells tend to react more extremely to the commotion / surprise. Again, put simply, it's harder to sustain a 25-beat rhythm when there isn't one or when it's just been 'lost'.

**4-star replies:**

R12. These accumulate – I like SJL's analogy of how in football two medium fouls can add up to a yellow card

R25. Important particularly when the competition standard overall is high.

R27. Ditto above.... Faults are by definition on a continuous scale. I appreciate/agree that it makes them easier to categories/compare by boxing them into discrete groups, but it should be remembered that this is purely to help us and does not reflect the reality.

<b>Persistent errors from a particular bell or group of bells</b>	*	*	**	*****	*****
---	---	---	----	-------	-------

**5-star replies:**

R17. Key information

R24. See earlier comment about judging 'persistent error'.

R25. This relates to my comment above in red about consistency across several changes – touches contain patches of good and not so good ringing.

**4-star replies:**

R12. Depends on the magnitude and how consistently/deliberately they're done within the group of bells.

R13. Marked 4 as it doesn't matter who makes the errors

R28. Comments from judges regarding certain bells would be useful but I don't think individual bells should be singled out. It might be better to say something like "one or two bells were struggling with odd-struckness" for example. If teams want to know which bells in particular, that could be something that the judges may provide to the team captains.

**3-star replies:**

R14. Makes no difference to what is good ringing, but identification of this can help to produce it.

R19. Not that such errors don't matter – they will be picked up by other metrics.

**2-star replies:**

R20. Not so important whether it's coming from the same bells.

**1-star replies:**

R27. Clearly this would be useful/important information for any individual band and will influence the overall quality as measured by eg pure SD, however I strongly believe that this should not be given any greater weight than the individual errors. I don't currently believe in itself (ie persistent errors vs the same overall errors spread equally through the band) will influence the overall quality of the sound produced and on top of that, any use in judging/determining quality in competition could be seen as particularly divisive.

***Persistent errors in certain parts of the change***

\*\*\*\*\*

\*\*\*\*

\*\*\*\*\*

**5-star replies:**

R17. Key information

R18. If there's a problem leading, or when doing places, it's good to be able to spot that

R24. See earlier comment about judging 'persistent error'.

R25. See box above.

**4-star replies:**

R12. Related to the previous answer. Not clear how to compare entirely rhythmical ringing with nearly always rhythmical ringing with deliberate variations for musical effect

R13. ... or where they are in the row (with the exception of leading but that's already covered)

**3-star replies:**

R14. Makes no difference to what is good ringing, but identification of this can help to produce it.

R20. Occurrences in the same part of a change is probably more discernible.

R27. Similarly to the above – likely to be very valuable/important to the band in terms of improving the ringing, but I don't believe it should be taken as being of any greater importance in defining the quality of the ringing *per se*.

***How well the bells (or groups of bells) work together as a team***

\*\*\*\*\*

\*\*\*\*\*

**5-star replies:**

R14. This is key to producing good ringing!

R18. Really important that you don't get multiple approaches within a band

R24. This is important because it can help you 'stay on track' through imprecision and error in different parts of the change. So, if the treble is a 'wonderful ringer', they will still be exactly 7 or 8 beats off the front as you go to do 7/8 places with them. They will have ignored the slow or quick leading at the beginning of the row and sustained the underlying rhythm. If you still have that rhythm in your head, they effectively 'confirm' that to you and that sustains you for (say) another 5 or 6 blows, whether there is minor imperfection elsewhere in the change or not. This is a really tricky area and one I think we benefit from a lot. If you ring together regularly and are a proper Sunday service band, then you can 'just know' (through osmosis) where your course bells or buddy bells or adjacent bells are going to be (when they are in the right place). We notice tiny differences even when we swap one 'wonderful' ringer for another 'wonderful' ringer in some of our 'units'. The weakest member of our team even notices this – and was quite surprised by this – it's a genuine intangible – perhaps it's the 15ms differences? Perhaps it's a different type of randomness in the equation?

R28. Comments from judges regarding how teams work together to produce good ringing is really important for morale. It brings teams together.

**4-star replies:**

R20. Fairly important, without this the feature changes e.g. roll ups are going to be uneven and disrupted, which will be unpleasant.

R25. Particularly at important points like big and little bell roll ups

**Non-marked replies:**

R27. How do you measure/define this? It sounds good in theory, but I remain to be convinced that any objective measure of this is available that could enable us to draw a view as to whether it could realistically be used to define good ringing. Happy to comment further on any specific measure that might define this. Similarly, in common with many of my comments above, whilst it may be useful/important to the individual band in informing what can be done to improve the ringing (cause and effect), we should not lose sight of the fact that the important measure should be based entirely/objectively on the sound produced.

<b>How well large and smaller bells work together</b>				*****	*****
<p><b>5-star replies:</b>  R12. There's a conflict here between sticking rigidly to the rhythm and adjusting to make the ringing locally good</p> <p>R14. No more important than how well similar groups work together, but without this you can't produce good ringing.</p> <p>R24. This is simply how a lot of bands accumulate a lot of faults / error / inconsistency. If you sustain the 25-beat underlying rhythm, this won't happen much.</p> <p>R28. As above.</p> <p><b>4-star replies:</b>  R20. Small bells clipping larger bells can be particularly unpleasant.</p> <p>R23. Although it may sometimes be more difficult to achieve precision in the interactions between heavier and lighter bells, I don't necessarily think this should be privileged when considering what constitutes good ringing.</p> <p><b>Non-marked replies:</b>  R27. Ditto above.</p>					
<b>Whether a particular bell or group of bells has been rung to a much higher standard than (most) other teams (e.g. specific odd-struckness or audibility challenges have been overcome)</b>	****	**	****	*****	*
<p><b>5-star replies:</b>  R24. If any of these things happen, then you will (and should) score more highly. I believe that the audibility challenges should be eliminated by only holding the Contest at towers with excellent acoustics. I do not understand why some people seem to prefer not to hold it at such a venue.</p> <p><b>4-star replies:</b>  R17. Useful info to acknowledge outstanding individual performances in a less successful team</p>					

R27. Worthy of further consideration. Could(/should?) certainly be part of subjective/artistic consideration by individual judges – BUT, magnitude of any importance in relation to the quality (relative to other factors – e.g pure SD) will by definition be (rightly) subjective (ie should NOT be defined by a single desired numerical value, and at best non-linear). I appreciate this causes issues in terms of “how do teams know what they are aiming for if it is not defined”, but as above for major/minor faults, the ultimate aim should be clear, it is purely the method/ranking of failures to reach perfection that is subjective....

R28. As above.

**3-star replies:**

R12. Shows excellence, but should be reflected in the results anyway

**2-star replies:**

R13. It’s a team game. The most destructive member of bands I have rung with has tended to revel in how they have the best SD.

R23. Might this variance be accounted for by the other areas described above rather than being considered an independent predictor?

**1-star replies:**

R14. While this may be important for band placing, improving ringing, or evaluating individual performance, it has no bearing on good ringing, which is a collective endeavour.

R15. Useful for judges to point out in competitions I guess, but those that already know how to strike well don’t need to be told – and those that don’t could almost certainly live without the comparison.

R20. It’s only the whole band performance that matters.

R22. This is not a metric for good ringing

<b>Quality over time – how quickly the touch settles and whether it continually improves</b>	**		*****	*****	
--	----	--	-------	-------	--

**4-star replies:**

R12. Key – it’s natural to like a piece which finishes well compared to one which fades away. However, not clear how these should be judged!

R25. Important, but I think a good a good piece of ringing needs to be good at the outset.

R28. Comments like these are encouraging to the team.

**3-star replies:**

R14. Not important as a metric in itself, but of course has a large bearing on the overall quality of a piece of ringing.

R17. We find this lead by lead analysis useful

R23. The salience of this point perhaps depends on whether a band has been able to practice on the bells? If there's been no practice, then I think it's reasonable to assess how quickly a band settles and whether their ringing improves over time.

R27. As with many of the comments – this will likely impact considerably on other measures (e.g. SD) and as such will be very important for individual ringers/bands in *effect* on the ringing (more so than defining the overall quality of the ringing itself). There may be some merit in consideration however: Are longer periods of fairly good ringing better than short periods of very good ringing if all else is equal?... this is very subjective (and non-linear) and in my view worthy of consideration, but likely to be something that [should] remain[s] personal choice (open to interpretation by individual judge/judging teams?) rather than a fixed metric/definition. Any additional information in terms of how this varies in any individual touch that can be made available to inform that decision would clearly be a good thing however.

**1-star replies:**

R20. I don't think duration in the piece has any bearing on whether a fault should be valued differently.

R24. Continual improvement is highly irrelevant – if it's still improving after 13 minutes, then it must have been 'terrible' to start with. How quickly the touch settles is really a measure of how quickly consensus over the 25 bell beat is established; it then has to be sustained. They are different skills / phases – but both need to be mastered quickly in order to be worthy winners. It is possible to hit a real 'purple patch' but realistically this almost never happens in a pressurised 13 minutes on unfamiliar bells (3 times in 26 years). Finishing well or on a high is something judges often comment on – in my mind, this is a consolation prize. This is one of the areas where Hawkear is effective because it is consistent throughout.

<b><i>How well bands cope with fault / error in a change (do they manage to quickly or immediately sustain the underlying rhythm?)</i></b>	*	*	*****	*****	**
--	---	---	-------	-------	----

**5-star replies:**

R13. Crucial

R24. Yes – this is pretty much the whole game.

**4-star replies:**

R12. Very important, but mistakes will very likely limit how well a band in this situation can do

R28. As above.

**3-star replies:**

R14. Not important as a metric in itself, but a band less able to cope with errors will consequently produce less good ringing.

**2-star replies:**

R27. Whilst important in itself, this should be picked up by other measures of the quality of ringing (quick recovery will lead to lower overall error in any case), so shouldn't in itself be important in defining the quality of ringing.

**1-star replies:**

R20. It's nice and may be helpful commentary, may even be a good indicator, but I don't see how this can be a metric in any absolute way.

**Non-marked reply:**

R23. Perhaps this variance be accounted for by the other areas described above?

**How well a particular row is rung (e.g. queens, or backgrounds)**

\*\*\*

\*\*\*

\*\*\*\*\*

\*

\*

**5-star replies:**

R20. I think this is vital and is totally missed from current algorithms. I suggest a new metric on this. Evenly struck feature rows and rollups should be credited, and failure to get this right should be penalised even if overall consistency is good. I suggest these rows are judged in isolation too – bells shouldn't be negatively scored for the feature row if they are slightly quicker/ slower than other changes, if the key row itself is even. An uneven feature row is the most audible, unpleasant and disappointing way for a fault to be experienced and therefore should be valued highly, not evenly, to all other changes. I would like to see some algorithmic tools made available for judges to score this, so that they're not left to judge this totally subjectively (and in cases, in contradiction to the hawkear outputs).

**4-star replies:**

R28. As above. It creates a positive feeling.

**3-star replies:**

R12. Part of the 'musical performance' aspect. It depends how the music builds up to it – Queens coming out at the first row of a six sounds different to the end of a six



R13. While it's nice to hear Queens come up perfectly, it's only one change out of 290 or so.

R14. Every row matters - significant ones can be a good opportunity to regroup (or hesitate!) but aren't much more important than a consistently good rhythm.

R17. Arguably it's just one row but inevitably they stand out to the listener

R25. One change among many. All changes are equally important although the human ear naturally appreciates certain ones more than others.

**2-star replies:**

R27. Probably unimportant. May be worthy of some minor consideration (laymans perspective of sound may well pick up on e.g. rollups more, such that this becomes more important subconsciously – should we be encouraging this focus further by emphasising it: I am not so sure!).

**1-star replies:**

R23. I don't think certain changes should be given privilege over others when determining whether a touch has been well rung or not.

R24. Answered above.

**Extra metric: R25. *Consistency for periods of ringing across several (say 10 +) changes.*** 5 out of 5. I think it's important to look at sections of ringing across several changes. If a touch contains errors but also delivers a patch of good settled consistent ringing it will I think, be thought better than a touch with fewer error but without such periods.

\* HawkEar, of course, treats all bells equally, but we are trying to understand whether the tenor, treble, or indeed any other bell is subconsciously perceived as more important than another to the ringer or human judge.

Reply to this comment: R24. Yes – well it's only really more important for its impact on the other human ringing live-time!

Reply to this comment: R27. My personal opinion here is that these are 'cause and effect' factors – consistency in the majority will have a significant impact on the individual accuracy of all/any bells and may have more physical importance in the band, but this should not be confused with specific effect/importance alone on the perception of quality. Also noting of course that the answers above will (for all respondents) be based on conscious perception and not reflective of the stated desire to understand subconscious perception.

## 2. How would you measure “overall impression” of a piece of ringing?

R12. It’s mostly determined by the effect of how rhythmical the ringing is i.e. how consistent the interbell gaps are, along with how the errors and deviations from this framework are sprinkled in

R13. By definition you cannot measure an impression! You can listen (and re-listen) to pieces of ringing and then rank them in your personal order of preference in some way. You might even assign it some form of score but that will inevitably be subjective and qualitative. Are there any data comparing a panel of judges individual “overall impression” scores with Strikeometer results? If not it’s an analysis that should be done.

R14. With difficulty! How much you sit back in your chair, or lean forward waiting to hear what the next row brings? A combination of how long you have to wait for the next fault, and the magnitude of the faults. It is certainly easier to rank (I enjoyed A more than B) than measure (A was 67% good).

R15. Ranked in order of how much the judges liked them, like other musical competitions.

R16. Number of mistakes, faults, consistency of striking over the piece of ringing.

R17. Using a percentage - it’s best understood by the largest number of people.

R18. By ear... There are several things which I’d do... 1) Has it got a good rhythm – ie. All bells are ringing to a similar plan; 2) Is it a nice speed? Not too slow, not too fast for the bells; 3) Trippy or mainly accurate; 4) How is the leading?

R19. Is it nice to listen to? Would I like to take part in ringing like that? Trying to quantify this I think that how good are the good / average bits is more important than how bad are the worst few bits.

R20. I’d put some Key Performance Indicators on a dashboard for the judges, and then tie up with subjective comments. My KPIs would be: 1) Standard deviation, speed adjusted; 2) SD, speed adjusted, with probably method/ transposition error excluded; 3) Fault count, > 50ms; 4) Special row interval SD; 5) Special row fault count > 30ms.

R21. By a percentage mark

R22. I don’t think I could. If you know how to listen you can tell how good the overall impression is.

R23. Listening to a piece of music will trigger impressions in us all which will depend on more than just the degree to which the bells are consistently equally spaced. Impressions may depend on the listener’s history – for example, a speed of ringing that is consistent with the speed the listener is familiar with or believes the bells should be rung at might trigger a more positive impression. I’m not sure this could be measured in any other way than measuring the listener’s level of satisfaction with

the ringing, but this probably can't be measured objectively over the course of a full day of judging. So, while feedback could be provided regarding impression and satisfaction, I'm not sure this could be objectively measured?

R25. Overall impression seems to me to be in the gift of the audience. While computers can analyse, only people can form an impression. So measure by asking people to listen and rate the ringing as a whole – ideally not by comparing to other pieces, although this is surely not possible in the context of a competition.

R26. Front, middle and back bells all in agreement over speed with the back bells generating a good structure for others to follow. Consistent leading and changes remaining relatively consistent in length, at least so far you can't easily hear changes being strung out etc. Good speed for the bells – i.e. close to average peal speed.

R27. I don't have an answer to this. Other than "overall impression" is clearly (and must by definition be) a vague and subjective measure, undefined by specific metric, so if it is considered that this is worthy of use in *defining (marking?)* good ringing, the only sensible way to do this is to leave it to the subjective/individual view of the person measuring it (judge/judging team). Definitely worthy as comment, but possibly difficult to include in precise definition or ranking comparison.

R28. As a judge, I think it's only possible if you have a great deal of experience in ringing on twelve and perhaps taking part in. Have a couple of the judges marking faults and another judge just listening to the ringing and getting an overall impression. This way they can write notes during the piece and not try and do both at the same time.

### **3. How would you define a touch's "compass"? And is it important?**

R12. Very important – it's the combination of the interbell gap and handstroke gap, and how these change or don't change from row to row

R13. It's a poorly defined term in relation to striking. Looking at the OED it's not easy to find any of the (non-obsolete) definitions which it easily relates to! My assumption is that it is used as way of describing the overall time envelope within which individual rows are rung and how stable this is across the whole touch. My mental image of this is of a flock of birds (such as a starling murmuration) and how they flex and move relative to one another. Assuming I am correct, it's important but we really do need to be clear about what we mean with any term when describing ringing and good striking.

R14. The overall framework made up of the rhythm set by the back bells, the leading, and the placing of the bells at the back of the change. It is probably easier to identify what deviates from the compass than what it is. Is the leading sometimes rushed? Does the back of the change string out? Perhaps the compass is the framework on which these judgements can be based.

R15. I wouldn't because I've never understood what compass means. It therefore has no importance to me.

R16. How the quality of ringing evolves over time. In a striking contest, it is important as to how quickly a piece of ringing settles down.

R17. Not a term I ever use.

R18. I wouldn't. It's a phrase that some people use that I have never heard anyone define, so it's a bit Mornington Crescent in my book!

R20. Probably by the tenor ringer's speed, and maybe this relative to overall speed. I don't think it's that important as a metric of how nice a piece of ringing sounds but may be valuable/ important to understand what may be driving imperfection.

R21. Consistency of speed and rhythm

R22. Consistency of length of change and it is important

R23. I'm afraid I don't know what a touch's compass refers to.

R25. Consistency of ringing speed and accuracy. Hugely important.

R26. In general, a piece of ringing with a strong compass will have a consistent speed with front, middle and back bells in agreement. That isn't to say that a piece of ringing can't speed up or slow down and not remain accurate, but in practice doing so is likely to pick up audible errors. So, yes, compass is important but not necessarily *the* factor that, in itself, determines a good piece of ringing.

R27. Unsure on writing down a single definition – I suspect it is pretty undefinable. Based on my current understanding, I suspect it would best be made up a number of individual factors based on rhythm within and between rows, and likely *correlates* well with (is not independent from) measures of quality determined purely by SD. Is it important – certainly not unimportant, and very interesting, but I am not convinced it is yet well enough understood/defined to either use in *defining* what is good, nor sufficiently independent of other factors to be any more important as a measure in itself.

R28. Not sure what this means.

**4. If you were comparing a quick piece of ringing to a slow piece of ringing, how would you discriminate between the two if both were very accurate?**

R12. By which sounds better for the bells. For bells with less hum, fast often sounds better, and for bells with lots of hum, slower is better

R13. It's difficult and to some extent would depend upon the external acoustics of the ring of bells if listening / judging from outside. Generally speaking a well-controlled accurate but faster piece of ringing is more pleasing than a slower equally accurate and well controlled piece but sometimes on deeper toned bells slower can sound better.

R14. Does one sound rushed, or the other sound ponderous? It is subjective and comes down to impression - as in comments like "falling over each other" or "tapping along nicely".

R15. I wouldn't. If they were both equally accurate I would give them an identical score.

R16. I don't think there is a need to – if both are equally accurate, it shouldn't matter as to the speed.

R17. This is where I'd expect the minute accuracy of HawkEar judging to be utilised. Personal preferences for quick or slow ringing shouldn't influence results at this level. I also wouldn't object to a tie if two teams were really that close.

R18. I'd consider the overall speed for the bells and whether it was turgidly slow or too quick and frenetic.

R19. I don't think that the precise scores according to particular metrics should be used to discriminate: my experiments with HawkEar have shown that precisely where in a recording transcription starts could potentially cause the order of teams according to a particular metric to change. It would be better to have a range of metrics, including an "overall impression" score to separate bands that come close, or have conflicting scores.

R20. I think this is very hard. Speed adjusted SD/ fault measures are important. I understand the algorithms do penalise slower ringing more. I wonder if the speed adjusted metrics can be developed to compensate more for the optimal strike position being statistically harder to find

R21. I wouldn't

R22. I wouldn't

R23. I think this is a thought-provoking question. I think there is probably a positive association between more impressive ringing and quicker pace, but only up to a certain point. Perhaps this positive association between speed and quality plateaus when the sense of rhythm one achieves by noticing and being able to discriminate 12 bells in a change starts to fade? A bit like when background hum inhibits one's ability to distinctly hear each individual bell in a change and therefore one's ability to sense rhythm is impaired. At a certain faster pace the same detrimental effect perhaps starts to occur?

R25. Difficult. I'm afraid it would come down to which I enjoyed listening to more.

R26. Unless we're talking two extremes, I'm not sure I would deliberately discriminate. If two pieces of ringing accrued the same number of faults or had virtually identical SD and needed separating, it is likely I would discriminate in favour of the piece of ringing that 'sounded' or felt like a better fit for the bells. Examples of extremes that were not pleasing to listen to would include the ASCY touch at Aston in 2016.

R27. I would want to look at both *actual* error and *relative* error and take a view, probably with other (possibly subjective) measures – e.g. see above in relation to magnitude of what constitutes major vs minor errors. I don't believe that there is a specific answer (i.e. to whether either is more important/accurate), nor that one method should be adopted. Whilst I understand/agree that this is currently often seen as a weakness (and can lead to unhappiness with individual results) this is one area where individual judge subjectivity should be maintained as whilst individuals will have their own views, these are subjective and we should welcome the range.

R28. Number of faults whether they be minor or major. Speed should not make a difference if it is well struck. Overall impression might be a way as fast ringing can sometimes be described as being “rushed”.

**5. When looking at HawkEar output, what do you find most useful? And what would you like to see that’s not there?**

R12. The main Touch Viewer page. Within this, it should be easier to jump to the biggest errors for a particular bell. Adding filtering (Excel pivot-table style not Kalman filter) to give metrics per piece of work - give an average and sd for e.g. all the down dodges

R13. It depends on context. If reviewing performance of our band then I use the individual bell error tables to look for outliers or outlier pairs and then try correlating this with good and less good parts of the touch. Perhaps one thing that should be developed is a visual representation of compass of ringing – once defined!

R14. The visual representation of the touch, which is a useful prompt to use while listening back to a piece of ringing. Other measures are useful, but transcription errors or other factors (a bell seeming slow because the previous bell rushes a lead, for example) can influence them.

R15. I tend to look at the ‘graph’ bit to see how badly I’ve done and if I’m feeling paranoid, my SD score too. I cannot make head nor tail of the rest of it. It would help if there was more (any?) discussion about how much the brain automatically tries to compensate for the impact that other badly-struck bells in the early part of a change can have on where they place their bell in the rest of the change.

R16. Number of errors over 50 ms, individual bell error histograms, grid of changes with colouring and late/early bars for individual bells.

R17. We use the Band Striking Quality and Errors over 50ms for headline feedback for the group to measure progress, then the Individual equivalents to guide individual feedback alongside the patterns of hand or backstroke inaccuracy. We analyse by lead to look for areas which consistently need improvement and then look for groups of bells unsuccessfully working together.

R18. I’ve updated it and xxxx uses a different version... (happy to show you!).

R19. For the band as a whole: Striking summary. For individual bells: Error histograms and Bell error tables. What’s not there: measures based on median and percentiles rather than mean and standard deviation; measurement of mean absolute error. (Rather more difficult) Measures based on the type of work being done, e.g. is there any difference in the accuracy of hunting and dodging?

R20. Depends on purpose: 1) If I’m trying to select individuals for a team, the individual SD measures; 2) If I’m trying to fix structural issues in the ringing, the bell-in-position data is very good for picking up tendencies that are usually transferable across towers and often individuals; 3) If I’m judging a piece overall, my first metric is overall SD speed adjusted error eliminated. Second metric is fault count > 50ms. There needs to be a metric for in-change evenness of feature rows.

R21. Whether there are any consistent faults being incurred. Nothing.

R22. I don't

R23. Most useful: 1) Average individual error; 2) Errors over 50ms; 3) Individual and band striking quality on the stars scale; 4) I find the use of colours to highlight slow/quick striking very helpful. Might it be useful to have a bit more information to interpret the significance of numbers? For example, I understand what 'random error at hand' and 'handstroke uplift' means but I don't know how to interpret the significance of my scores. For example, if I achieve 26.9ms random error at hand, can I expect myself to improve this and if so, how?

R25. All useful. Perhaps some clearer indication of which parts of the touch (i.e. which leads maybe) generated the most errors. We often practise the first or last leads because we want a strong start and a strong finish, but those leads might be our best ones while errors are cropping up elsewhere.

R26. Instinctively, I first look at the SD. However, in terms of aiding my own ringing I'd look at mean error by bell hs and bs and then following my bell through the touch combined with studying the bell error tables to see where there is obvious room for improvement.

R27. Undecided currently. Certainly the output is interesting. Is it useful – I would like to say yes, but I am really not sure how I actually make use of it on an individual basis (though this will similarly be answered differently depending on whether we are talking about making use of it to try to improve my/band ringing or in determining/comparing the 'objective' quality of any individual piece(s) of ringing.

R28. The star rating of the touch as a whole as well as the star rating of individual bells.

## **6. Would you like to see fault counts or standard deviations?**

R12. Both – see answer to 2). They combine to give the overall impression.

R13. Both! But also be very transparent at how final percentage scores are reached and how they relate to these two measures ... and any other measure whose definition or derivation isn't published – like the "Mark" which determines placings.

R14. Both for information, but fault counts (possibly in whole and half faults) as an overall mark.

R15. Neither! I'd use HE as a tool for rehearsals i.e. similar to those practising scales with a metronome or clic track when learning, and then to iron out persistent faults thereafter. But using it for 'performances' (i.e. Sunday service and competitions) is akin to the tail wagging the dog IMO.

R16. Both.

R17. It's important to reduce jargon and technical terms as much as possible in every area of Hawkear reporting as it can be a deterrent to use and add a burden to those who have to explain the terms. Even the recent Zoom session was reported to be over the heads of some of those who logged in because of the jargon used. Fault counts need no explanation, standard deviations do.

R18. Both. I don't think there is one metric to rule them all.

R19. Standard deviations are more objective, while faults rely on a somewhat arbitrary threshold, but both have drawbacks:

- large errors have a disproportionate contribution to standard deviation
- quantization errors mean that using a fixed threshold for measuring faults is inaccurate

If measuring faults I would prefer to see a sliding scale between 0 and 1, e.g. 40 ms or less scores 0 and 60 ms or more scores 1, with a linear scale in between – this removes the cliff-edge and quantization errors should cancel out.

R20. Both are very valuable. For improving ringing SD is probably most helpful. For judging ringing, I think fault counts, and how those faults are scored, is most important.

R21. Fault counts

R22. It doesn't matter because neither is an absolute measure

R23. I won't answer this because although I understand standard deviations I don't fully understand how they can be applied in this context.

R25. Fault counts mean more to me.

R26. Fault counts.

R27. As above – I think it would be beneficial to have greater ability to view both, along with play with the thresholds for faults and/or non-parametric measures (ranking).

R28. Fault counts.

## **7. How do you assess good ringing?**

R12. By how relentless it is, and how well the band builds and then maintains the framework

R13. What I look for is a solid, positive rhythm with the trebles and tenors ringing together seamlessly and everyone in between fitting in perfectly too. In terms of how it's assessed I welcome the consistency and reproducibility that HawkEar / Strikeometer has brought to the contest. However, I remain concerned that it is a black box in many ways and that some of the technical limitations as exposed in some eliminators and finals (Aston as one example) have not been fully discussed and some apparent anomalies (eg 3<sup>rd</sup> place at Sheffield seems hard to reconcile with the hard data) are never explained. More transparency please!



R14. I see good ringing as a lack of surprise at when the next bell sounds. Well struck, consistent ringing gives no surprise at all - you can predict where the next bell is. Slight inaccuracies give constant low level surprise at whether a bell is going to be quick or slow, and a mistake or major striking error gives a large amount of surprise. The quicker a touch settles, the more consistent and positive the rhythm, and the quicker any mistakes are recovered from, the more expectations are set and met. Of course, some aspects of predicting the next blow can't be factored into this. If the tenor is always covering in N+1th place at backstroke, that is predictable but still not conducive to good ringing.

R15. I work out who the good ringers are in every touch and try to benchmark my ringing around them so I don't clash with them as much the next time. Please don't get me wrong: I see HE as an incredibly useful tool to help learners really think about accuracy. Without using it for the 12 bell I doubt it would have developed in the way that it has and I really take my hat off to those that have put in a huge amount of time and effort in over the years. It does have its downsides though. For me slightly playing around with the rhythm is the nearest equivalent to a woodwind player slightly sharpening the leading note so that it resolves better onto the tonic i.e. it's musically satisfying. It could be characterised as 'doing a David House' (where really skilled ringers predict the music of the next change then flick their wrists/bend their elbows/open up their shoulder blades to really make the most of the change) but this would now be counted as a fault by Hawkear. That's really sad. I accept that this always was a vanishingly rare skill, but it spoils the enjoyment of ringing for me.

R16. Number of mistakes, faults, consistency of striking across each member of the band and over the piece of ringing.

R17. My priority is how pleasant it was to listen to. Was there a flowing rhythm, a clear beat and accurate striking providing that. An extremely accurate ringing friend said they wanted to feel they could dance to it.

R18. See Q2... few faults, nice back bell rhythm, accurate front bells, runs well struck, leading consistent, not too slow, not too fast.

R19. For the most part, good ringing will mean good HawkEar scores. What I think is important is that the use of HawkEar should be transparent: publish the metrics that are being used to assess the ringing, and most importantly, explain why judging has gone against the HawkEar results according to these metrics. If there is a good reason not to go with Hawkear, come up with new metrics that do capture the difference.

R20. With hawkear outputs – as indicated in response to Q5. With human ear – I care much more about the sense of rhythm and the evenness of feature changes and the rows around them. Absolute audible faults, particularly clashes involving the bigger bells I also find more unpleasant e.g. if T hits E, or 1 hits T.

R21. Is good to listen to, consistent leading, good rhythm being set by the tenor and the rest of the bells following that rhythm.

R22. I listen

R23. I think minimal audible faults or clips with a consistent pace and consistent gaps between bells, which means consistent separation is achieved within a change that starts and ends at the same point as the other changes i.e. 'inside the box'.

R25. A combination of all of the things in this survey, but certainly the perception of the sound by a human audience should remain the most important factor.

R26. Primarily, we felt there should be a greater reliance on the 'fault or error counting' and 'team' areas in terms of the metrics judges should most heavily rely on. There will, of course, be correlation between this and what Hawkear says, but what the human ear can pick up when judging should take precedent. We have to accept there is a degree of subjectivity here, and that is how it should be.

R27. Subjectively (!), with a strong focus on SD, but some interaction between consideration of SD of individual bells/place vs overall SD and associated ranking (non-parametric approach). I would be interested to see how much difference any of these different variable would actually result in changes to the ranking of individual pieces of ringing: if they never/rarely do, then I would suggest we don't need to worry about it at all. If they frequently/sometimes do, I would suggest that this is more likely reflective of subjectivity (which should be welcomed/incorporated through individual preferences rather than cut out by taking us down the route of focussing on a single set of metrics).

R28. In terms of judging a competition, I grade faults from good to bad as: 0 – perfect blows; 0.25 – slight hesitation or slightly quick blow; 0.5 – clip i.e., a near miss; 1 - clash i.e., 2 bells ringing together at the same time. I also try and get an overall impression of the ringing if I can. Some touches can be almost perfect ringing but sound "mechanical" whereas other touches "flow".

#### **Email replies:**

R13. I'm glad to see the work being done but still have a few reservations as you will see. Perhaps the two most important limitations in the current system could be expressed as two issues from recent contests: (1) "I had to deliberately ring my bell too close for the strikeometer" (Aston 2016, Southwark 2017 - different members of the same team) i.e. Shouldn't we judge what the ringers are hearing. (2) "How did they come third with those scores?" (Results of Sheffield contest 2023) i.e. Let's have a bit more transparency about how the data become the final results!

R17. As a team captain with absolutely no understanding of the maths behind Hawkear, but who uses it regularly to feed back information to the team, I'd like to plead for as little jargon to be used as possible in communications about it. I'm sure it's done in complete innocence, but for those of us totally outside this area even using the word metric to head a column rather than something in more general use like criteria is intimidating. If you want a "Hawkear for Dummies" candidate, I'm your man!

R20. My main thing that I think is missing is an algorithmic representation of how well feature rows are performed e.g. a metric for in-change evenness of feature rows.

R22. I don't want to be negative but I actually found quite a few of the questions rather simplistic, given the preamble in your original email. However I'm sure the Committee will be able to extract some useful information from the replies received.

R26. It was difficult to clarify our thoughts, but I think it's fair to say the general weight of opinion was in favour of the more traditional methods of defining good ringing.

R28 replied by email rather than completing the survey questions: It would be interesting to manufacture two pieces of ringing electronically which achieved the same Hawkear score but in different ways. Or one piece of ringing that had everything perfect but a sprinkling of errors over 30ms which judges pick up, and then another piece

which Hawkear might fault more heavily because there is lots of inconsistency but no single error is over say 25ms and hence could not be detected. Which would sound better or worse? Which should we think is the better piece of ringing? Then there are some things which are almost self fulfilling, like the handstroke leading. I don't think I notice inconsistent handstroke leading per se, but I would spot the quick ones and the slow ones. But if you have inconsistent leading then it's going to have such an impact on the rest of the ringing that the leading bit is almost insignificant. People say they can tell the Birmingham band's ringing but is that because it has a style or just because it is (often) just better? My team have been criticised for ringing to the strikeometer, which is unbelievably difficult so ought to be commended, but that doesn't necessarily make the ringing sound nicer to the listener. The xxx final was an extreme example when we adjusted my striking in particular to suit how the microphone was positioned and where the judges were going to be. The Devon ringers talk of the era of when the North Devon bands were unbelievably good – Eggbuckland would win a comp with virtually no faults, but that not everyone liked their ringing because it had no style. It seemed too clinical. Perhaps the human ear actually likes slight imperfection because then it appreciates the perfection even more? So that's a flavour of why I think it's difficult to answer those specific questions. I don't really know what it is about the ringing that makes it good other than that it is like a pond without any ripples in it.