

Commentary on Huovinen's "Varieties of Musicological Empiricism"

ERIC F. CLARKE
The University of Sheffield

ABSTRACT: Erkki Huovinen's "Varieties of Musicological Empiricism" provides a valuable analysis of some of the theoretical predicaments raised by pursuing an empirical musicology. But in this commentary, I argue for a less programmatic, and more pragmatic, approach to the term than he does. Empirical approaches in musicology have been around in one form or another for a long time, and the purpose of the label is less to identify a new 'brand' of musicology than to bring together a diversity of approaches that in different ways capitalise on the opportunities that data collection (in the broadest sense of the term) may provide. If programme is set aside in favour of pragmatism, and a looser relationship between theory and observation accepted, then empirical musicology can be a productive way to rub ideas up against a stimulatingly resistant world.

Submitted 2005 November 24; accepted 2005 December 5.

KEYWORDS: *Empirical, empiricist, pragmatic*

I will start with an anecdote: In a recent seminar, a student who had read an extract on empirical approaches in musicology and was considering what the word 'empirical' might mean, suggested that it had something to do with empires, the taking and controlling of a particular territory. This temporary confusion between 'imperial' and 'empirical' nicely illustrates what I would argue empirical musicology is *not*: it is not an attempt to lay claim to an area of musicology that is labelled or endorsed as 'empirical' (by contrast with an implied 'non-empirical' counterpart), and which is defined by the methods that it adopts and the aims that it pursues. More modestly (and perhaps benignly) the term serves to draw attention to the very varied ways in which empirical methods, understood rather inclusively, can be used to address or open up areas of musical research and understanding that may benefit from a more explicitly 'data-collecting' or observation-based approach. In that sense, the term 'empirical musicology' identifies a current within musicology that already exists (and has done so for quite a while).

Erkki Huovinen's paper "Varieties of Musicological Empiricism" is a valuable discussion of many of the theoretical questions (including central themes in the philosophy of science) that are raised by using the term 'empirical musicology', and since the paper touches on a number of points arising out of the book *Empirical Musicology. Aims, Methods, Prospects* (henceforth *EM*) edited by Nick Cook and myself (Clarke and Cook, 2004), it is from the perspective presented in that book that my commentary is largely organised. Huovinen's title already signals a different outlook from the one represented in *EM*: deliberately or not, Huovinen's use of the term 'musicological empiricism' puts the emphasis on empiricism as the approach, and musicology as the domain. I see the enterprise the other way around: it is musicology (again understood in an inclusive manner) that is the primary focus, with the book's purpose being to exemplify and discuss the varied ways in which empirical methods can be interesting and fruitful.

Simply documenting and taking stock of the empirical methods and perspectives that people have developed in various parts of the discipline reveals the diversity of both the methods and the subject matter that the term embraces. And of course this is only what people *have* done: the intention in *EM* was to gather together different varieties and examples of empirical research for at least two reasons: to provide a resource for anyone wanting to do empirical work in musicology, but who was unsure about how others had gone about doing it and what the opportunities (and stumbling blocks) might be; and to raise questions about what the aims of empirical work might be, and to stimulate suggestions and debate about how it could develop in the future. Because quite a bit of empirical work on music comes from, or is identified as, psychology of music, there is a danger that empirical methods in musicology are regarded as essentially psychological in orientation – that empirical musicology is more or less reducible to, or parasitic upon, the psychology of music. But it is important to recognise both that there is a lot of empirical work in musicology that is not psychological, and also that there are significant differences between the aims of psychology and musicology (even if it may

be difficult at times to be completely clear about what those differences are). As I have argued elsewhere (Clarke 2003), in very broad terms, psychology tends to be concerned with general principles (of perception, memory, action, thought), frequently stripped of their specific cultural and historical context, and often proposed as cross-cultural or even universal human attributes. An archetypal example might be George Miller's famous and much-cited paper (Miller 1956) in which he proposes the pervasive principle that human short-term memory (and the capacity to make absolute judgements within a perceptual continuum, such as saltiness, brightness, or loudness) is limited to seven plus or minus two independent items. The cultural and historical 'scaffolding' of different human societies may enable different ways in which people overcome this apparently rather narrow limitation (language, number systems, musical traditions), but the central psychological claim is one that argues for a universal human attribute. By contrast, musicology tends to be far more concerned with particular phenomena that are explicitly anchored in time and place: a study of the significance of birdsong in Messiaen's music *may* relate to more general questions of how nature is represented in music, but the primary objective is likely to be much more specific to this particular individual and his life and circumstances. One consequence of this (rather crudely sketched) difference of basic orientation between psychology and musicology is that musicologists become impatient with what they see as the generality (or blandness?) and lack of cultural or historical sensitivity of psychological work; and psychologists can be dismissive of what they regard as the unsubstantiated or insufficiently systematic claims of musicologists – based as they often are on detailed consideration of unique or isolated phenomena.

One of the purposes of *EM* was therefore to show that there are all kinds of empirical work in musicology that have no psychological component, and that empirical methods can be used in the service of a huge variety of broadly musicological aims. Musicology is not a scientific discipline, even if it incorporates approaches that make use of scientific methods of one sort or another. It encompasses various kinds of 'discourse' for framing, making sense of, and valuing music, and for that reason it should not be expected to adhere to scientific principles of consistency or verifiability/falsifiability. The desire to avoid the potential conflation of 'empirical' with 'psychological' is one reason for the possibly surprising omission (as Huovinen points out) of perceptual responses from the list of types of empirical data that appears on page 13 of *EM*. (Another is the oversight that may paradoxically result from something being so obvious that you fail to notice/remember it!)

Huovinen argues for a distinction between facts and data based on a qualitative/quantitative distinction – and here I take a different view. Qualitative methods in the social sciences have been increasingly accepted and acknowledged over the past 10-15 years, and rightly so. Qualitative methods can be very revealing in situations where quantitative approaches are impossible or inadequate. Huovinen seems to argue, however, that the material of a qualitative investigation should not be described as 'data' but as facts. I don't see the rationale for this: it is certainly true that quantitative and qualitative data are very different, and require or allow for very different kinds of analysis. But it has become routine to talk of qualitative *data*, and from an etymological point of view, as well as 'custom and practice', qualitative data such as the words spoken by an informant may be just as much the 'givens' (Latin *datum* = given) of a piece of research as are the scores on a rating scale or the timing values in a piano performance.

On the subject of qualitative and quantitative methods, Huovinen misunderstands my remarks about the relationship between data gathering and interpretation. Rather than *preferring* quantitative methods in which "the interpretative assumptions ... have become so deeply embedded as to be invisible" (Clarke 2004, 92), as Huovinen suggests, my intention was to do no more than point out that this was often the case – and actually to highlight the danger of that 'invisibility'. The conventional assumptions of standard quantitative methods (for example, the setting of 5% probability as the threshold for statistical significance) actually run the risk of appearing to naturalise their arbitrariness; whereas the case that new methods have to make for their own interpretative criteria makes it clearer that those principles indeed represent a deliberate and perhaps arbitrary decision rather than a state of nature. In that sense, qualitative methods can actually come across as more honest: far from holding the view "that it does not matter if our observations are theory-laden as long as there is enough consensus among the researchers so that we don't have to think about it" (Huovinen 2006, 6), I'd argue for the need to remain aware of the contingent quality of that theory-laden state.

If the reflexive character of qualitative methods potentially confers the advantage of a certain kind of explicit honesty, it also brings a potential problem that Huovinen identifies. As he observes, there has been a significant tendency in qualitative research for what is described as its 'reflexive' character to be claimed as an important positive attribute. The inevitable involvement of the researcher in his or her own research is, like the hidden assumptions in quantitative methods, something that should be acknowledged – but it is a different matter to make a virtue out of the interweaving, or even confusion, of the researcher's own experiences and attitudes with those of the informants with which he or she is working. As Huovinen points out, it is certainly easy for the two to get entangled, and arguably a complete separation is impossible to achieve: but as he points out, "this is no reason for an

empirical researcher to give up attempting to determine whether it is his or her own experience or the experience of the informants that makes up the primary empirical component of the research.” (Huovinen 2006, 8)

In general, Huovinen is more concerned to be systematic and classificatory in his approach than I am inclined to be, or would argue for. That may be down to a difference in intellectual style, but it also points to a possible difference in more general outlook. As mentioned at the start of this commentary, I don't see empirical musicology as a field to be defined, but rather as a collection of approaches to be used. In his remarks on the apparent omission of data from psychological experiments (see above), Huovinen states that “for the purpose of getting a firmer grip on the slippery field of empirical musicology, it would be useful to have a more well-grounded classificatory scheme for differentiating the various strands of empirical research from each other as well as from non-empirical research orientations.” (p. 5) The penultimate section of the article is then just such a discussion of a classification of ‘musicological empiricism’ into three strands: systematic empiricism, psychological empiricism, and historico-analytical empiricism. The classification is interesting as a formal clarification of the sometimes hidden assumptions that lie within different approaches, and it exposes some of the ways in which these distinctions might become unhelpfully blurred. But the emphasis here is closer towards classical hypothesis testing than I would want to go. At the end of the paper, Huovinen writes that ‘empirical musicology challenges us to identify and to discard those musical theories for which there is no “fact of the matter”.’ (p.12) I am more inclined to allow musical theories to continue to play a hermeneutic role, even when they fail to tie up with any specific ‘facts of the matter’, but to recognise that discrepancies between theory and observation can have a productive effect in making us look more carefully at both sides (or perhaps all corners) of situations that are nearly always more complex than they might appear. [1]

THEORY, OBSERVATION AND RESISTANCE

In a section entitled ‘The problem with theories’ Huovinen tackles the sometimes problematic relationship between theory and data, referring amongst other things to Cassirer's (1944) discussion of the theory-laden nature of observation. Without wanting to deny the significance of this insight, the reality is often that a far more provisional relationship between theory and observation is accepted in the practical circumstances of doing research. This can, of course, give the false impression of theoretical transparency, but it can also provide for a more informal and straightforward complementarity between theoretical and empirical components, a kind of tension that is never expected to reach closure. This may not amount to a very good or sustainable philosophy of science, but it is how a lot of investigation goes on: hunches become theories which are then examined for their adequacy, or sometimes get more rigorously tested, and lead in turn to modifications or abandonments of theories and new ones in their place, and so on.

I vividly remember the painfully ‘resistant’ quality of empirical data, and the eventually much more interesting view that this resistance led me to, in my own doctoral research. Inspired by a combination of what was at that time recent metrical theory (e.g. work by Yeston (1976) and Lerdahl and Jackendoff (1977)), and empirical work on piano playing by John Sloboda (which was eventually published in Sloboda (1983)), I set out to ‘demonstrate’ (as I was sure I would) the systematic relationship between rhythm, metre and expression. My firm belief was that if a rhythmically differentiated melodic sequence was rotated through a metrical cycle (i.e. started on each of the available beats of the bar), then pianists playing this set of melodies would produce performances in which the expressive properties of each note in the sequence directly reflected its metrical position. In other words, if the first quaver of the melody, in a 6/8 time signature, fell on the main downbeat it would show timing, dynamic and articulatory properties that directly reflected that metrical function; and when in subsequent rotations it fell on the second, third, fourth, fifth and final quavers of the bar, it would reflect each of those unique metrical positions. Armed with this appealingly simple (to me, at least!) hunch, I set about constructing a melody that would satisfy the needs of this study: it had to be rhythmically differentiated; it had to ‘work’ as a melody (I didn't want an arbitrary collection of random pitches); and it had to ‘work’ equally well in each of the metrical positions into which it was rotated. And here I encountered my first empirical resistance: as I should have realised, perhaps, meeting all three of these demands is not easy – even in a single line melody. What I witnessed as I struggled with this compositional task, was the way that notionally the same rhythmic attribute of the sequence dramatically changed its character as it ended up on a different beat of the bar, as did the melodic attributes. Intuitively, it seemed to me, this rotation was producing all kinds of unanticipated functional ‘mutations’ within the sequence itself. Still, I persevered and ended up with a sequence that had what I regarded as a sufficiently diffuse tonal character to allow the melody to make equal sense at each metrical position, and which had a rhythmic pattern that resulted in more or less the same total amount of syncopation in each rotation. I was ready to start collecting performance data.

I enlisted a graduate colleague of mine, who was an excellent pianist and improviser/composer, and having told him what I wanted him to do, placed the written materials in front of him and encouraged him to practice and get used to each of the six versions of this peculiar ‘non-tonal’ (as I described it to him) melody. I was a little taken aback when almost the first thing he did after playing it through a couple of times was to observe that it was a curious melody, basically in A flat major – promptly producing a rather convincing improvised harmonisation that did just that. After he’d done a little more practice, I duly collected my performance data from him, and subsequently from a number of other willing volunteers – and then encountered my second and more substantial empirical resistance. It was soon evident from the data that my belief in a neatly systematic relationship between metrical position and expressive attributes was premature, and that the picture was either a lot more complicated – or just completely different. And so began a longer process of sifting carefully through the data, looking at the possible dependencies and complementarities between different aspects of expression (timing, dynamics, articulation), thinking much more carefully about what the metrical rotations seemed to do to the tune itself, and considering how these unanticipated functional changes might not just drive the expressive features of the performance but might interact with them in various ways. The disappointment at not getting the result that I thought I wanted eventually gave way to a realisation that the situation was more interesting than I had imagined (it certainly gave me a lot more to write about!), and what I hope was a less naïve understanding of the relationship between structure and expression.

I started this commentary with an anecdote. The purpose of this closing anecdote is to illustrate the fruitful complementarity between theory, hunches, empirical investigation and even hermeneutics. Without rubbing my ideas up against hard reality (of both the compositional and data collection kind), I could easily have clung on uncritically to my firm belief. The resistance of my informal (compositional) and more formal (performance data) evidence forced me to think again – not simply to accept what might be thought of as the implacable ‘truth’ of the evidence, but to consider both the nature of the evidence *and* the claims of the theory. There are, of course, many possible explanations for the kind of disjunction between theory and data that I encountered, pointing towards the theory, the empirical methodology or the interpretation. In classically Popperian fashion, my data could be regarded as a falsification of the simple metrical theory that drove it. Or it could be that the materials (i.e. the rotated versions of the sequence) did not do what they were supposed to do: perhaps despite my efforts there was a systematic difference in the amount of syncopation in different versions. Or perhaps my participants and their approaches were just too heterogeneous – some simply playing what was in front of them without trying to find any sense in it, while others tried to find or impose some kind of shape and direction in the sequences. Or perhaps in analysing the data I was looking at separate dimensions (timing, dynamics, articulation) when I should have been looking at three-dimensional composites. The possible explanations of a prediction-data mismatch are always numerous, and it is out of a careful consideration of as many of those options as one can manage that new insights become possible. In the specific case that I’ve described here, there were sufficient questions about my own construction of the materials, the possible strategies of the participants, and the plausibility of the original predictions of the theory to justify careful thought about every aspect of what had turned out to be a much more complex situation than I had originally imagined. The longer-term consequence was a realisation that, however appealing it might seem, it was too simple to view performance expression as driven by structure.

To conclude, then, I am less attracted to musicological empiricism, if that means putting the empiricism first and the musicology afterwards, than I am to an empirical musicology that puts an understanding of music first and sees empirical methods as just one way to ensure against the complacency of apparently totalising theories. By contrast with Huovinen’s more intellectually rigorous analysis of the different varieties of musicological empiricism, my own perspective is a kind of pragmatic bricolage. By rubbing attractive ideas up against the hard and bumpy edges of empirical reality, there’s the chance both to understand that reality in a different way *and* to avoid getting too hidebound about our own pet theories. It’s not the only way to make progress, but it is one way.

END NOTE

[1] Though not always. Empirical methods can also be useful in stripping away layers of theoretical complexity, to show that a situation really is simpler than a tradition of theory might have made it out to be. Seen from one perspective, this is one of the consequences of David Huron’s analysis of the perceptual basis for the principles of voice-leading in Western classical music (Huron, 2001)

REFERENCES

- Cassirer, E. (1944). *An Essay on Man. An Introduction to a Philosophy of Human Culture*. New Haven: Yale University Press.
- Clarke, E. F. (2003). Music and Psychology. In M. Clayton, T. Herbert and R. Middleton (Ed.) *The Cultural Study of Music. A Critical Introduction*, pp. 113-123.
- Clarke, E. F. (2004). Empirical methods in the study of performance. In E. F. Clarke and N. Cook (Ed.), *Empirical Musicology. Aims, Methods Prospects*. New York: Oxford University Press, pp. 77-102.
- Clarke, E. F. and Cook, N. (2004). (Ed.) *Empirical Musicology. Aims, Methods Prospects*. New York: Oxford University Press.
- Huovinen, E. (2006). Varieties of musicological empiricism. *Empirical Musicology Review*, Vol. 1, No. 1.
- Huron, D. (2001). Tone and voice: a derivation of the rules of voice-leading from perceptual principles. *Music Perception*, Vol. 19, No. 1, 1-64.
- Lerdahl, F. and Jackendoff, R. (1977). Towards a formal theory of tonal music. *Journal of Music Theory*, Vol. 21, No. 1, 111-171.
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, Vol. 63, 81-97.
- Sloboda, J. A. (1983). The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology*, Vol. 35A, 377-396.
- Yeston, M. (1976). *The Stratification of Musical Rhythm*. New Haven: Yale University Press.