

Review of Noyce, Küssner and Sollich: Quantifying Shapes

DANIEL MÜLLENSIEFEN
Department of Psychology, Goldsmiths, University of London

ABSTRACT: This review discusses the original contributions and strengths of the target article by Noyce, Küssner and Sollich and also suggests a few points for consideration to maximize the usefulness of their analytical approach for the music psychology research community.

Submitted 2013 May 15; accepted 2013 June 14.

KEYWORDS: *audio-visual processing, shapes, contour, clustering, music drawing*

THE manuscript “Quantifying Shapes: Mathematical Techniques for Analysing Visual Representations of Sound and Music” proposes a novel and very interesting approach for gaining insight into the cross-modal processes that are linked to the perception of music. Specifically, the paper suggests several techniques to model spontaneous drawings in response to listening to single-line melodies and short audio excerpts. An interesting feature of the experimental task through which the data was gathered is that participants were not constrained in what or how they were drawing in response to the music they listened to. The data generated have a good level of ecological validity in the sense that these drawings might realistically represent how people draw (to) music in everyday contexts. On the other hand, this ecologically valid set of raw data also represents a big challenge for analysis because there are no standard techniques in the arsenal of music perception and cognition research that would be readily applicable to this type of data. As a consequence the authors have to find and adapt suitable analytical methods from the latest time-series and clustering literature to generate results that relate parameters of the acoustic input stimuli to the visual drawings captured in three dimensions (x- and y-coordinates on the drawing plane as well as pressure of the stylus used for drawing).

This review will first praise the strengths of this study and then consider some issues critically before ending on some points that should be stressed in order to maximize the usefulness of this study for a music cognition readership.

STRENGTHS OF THE TARGET ARTICLE

There are several aspects of this paper that make this study an interesting and valuable piece of research that adds to the existing literature. First of all, the sample size is relatively large for this type of experimental paradigm where the participants’ responses were constrained only minimally. Seventy-one participants drew 20 images each in response to audio files, and the resolution of the audio as well as of the image files is fairly high, sufficient to capture all relevant expressions. It is also noteworthy that the authors use a mix of artificial audio stimuli and excerpts from real recordings, which allows them to compare responses to full music with responses given to simplified musical stimuli which are analytically more tractable.

The experimental paradigm is ecologically valid and resembles how people would draw naturally in response to music. This high ecological validity is often sacrificed in experimental psychology studies in order to obtain simple and clean response measures which lend themselves to easy statistical analysis with common significance tests. The authors of this paper, on the other hand, make an effort to avoid this logic of experimental design (also often paraphrased as ‘I have a hammer and therefore my problem needs to be a nail’). Of course, the cost of this additional ecological validity is that, in order to make sense of the resulting messy data set at all, complex mathematics are necessary. Finally, it is worth mentioning that this study aims at modeling cross-modal processing of musical stimuli between domains including the time domain. It might not be the first study to model the functional and dynamic connections across two different modalities, but it certainly sticks out from the vast mainstream of cross-modal studies (e.g., literature on the effects of film music) where features of the objects in auditory or visual domain are often

reduced to single values or class labels and where a dynamic approach tracking perceptual changes over time is not within the scope of the research design. Still, there are a few issues that need to be considered in order make the approach presented in the target paper maximally useful for researchers in music psychology.

ISSUES FOR CRITICAL CONSIDERATION

For a better understanding of the range of the (presumably) very different drawings among the 71 participants, it would be very helpful to include examples from different figurative and analogous drawing styles and to discuss them very briefly in a qualitative manner. These more qualitative descriptions might be part of the other two papers referenced in the manuscript that are in press at the moment, but it would not hurt to include a few example drawings in this paper as well.

A second point concerns individual differences in drawing ability and expertise in expressing mental representations graphically. This type of expertise could either be collected with a suitable self-report instrument or even better by using an adequate test. Musical expertise was assessed and used as part of the analytical design, and musical expertise is certainly a relevant factor considering that the level of sophistication in the cognitive processing of musical stimuli is partly responsible for the drawn responses. However, expertise in the response modality might be equally important when trying to classify individuals purely on their drawings. Admittedly, it might be too late, now, to collect the relevant data and incorporate it in the analytical design of this study.

A third point, which is briefly addressed in the manuscript but could be exploited more in detail, concerns the references to the existing music psychology literature aiming to describe similar processes. On the one hand, there are a few studies out there that model responses to music over time using different and mainly simpler mathematical approaches; for example, Schubert's time series modeling (Schubert & Dunsmuir, 1999), functional data analysis as employed by Vines and colleagues (Vines, Nuzzo, & Levitin, 2005), or the music and body-gesture alignment pursued by the groups in Ghent, Oslo, and Jyväskylä. Researchers with a music psychology background who have followed this strand of the literature and already have some understanding of those earlier techniques would certainly be interested to know in what respects Gaussian processes and spectral clustering are similar and how they might complement or even go beyond what is possible with, for example, traditional time series analysis or functional data analysis. The same criticism applies to the research question and hypotheses that have driven this study. The motivations for this study appear to be three-fold: The desire to model free-form drawing in response to music, the curiosity to apply very advanced analysis techniques to this almost unconstrained response data, and the aim of comparing musicians and non-musicians on the drawing task. (It is not entirely clear, however, whether the focus of this comparison is statistical classification or rather a descriptive task). None of these motivations seems to be driven by perceptual theories or specific models from music cognition that could be checked or maybe even supported or rejected by the evidence gathered in this study. Tying the results of this study more closely to models and theories that are currently debated (for example the embodied music approach or the literature on gestures) would help to explore the implications of what is reported in the results and conclusions.

SUGGESTED POINTS FOR GREATER IMPACT

To summarize, and in order to suggest a few amendments that would help to make the approach outlined in the paper maximally useful to empirical researchers in the music psychology field, I suggest considering the following points.

The way the mathematical analysis is described in the paper is probably not very compatible with average level of technical knowledge that can be expected from researchers working in music psychology. Hence, to enable those researchers to adopt or even to consider Gaussian processes and spectral clustering for solving data analytic questions in their own research, it would be necessary to include non-technical descriptions of the basic principles underlying these techniques in the footnotes (in addition to the technical details that are already given there). Also, it might be a good idea to describe in more detail the specific usage of the MATLAB toolboxes that were employed, so that interested readers could use these descriptions as examples to follow.

I also think it would be very important to clarify with much more detail how this paper relates to previous literature from the music perception and cognition fields. Describing related models that also use multi-dimensional time series approaches (even if the data did not come from the visual modality but represented emotions or bodily movements), and discussing overlaps and differences, would help in

understanding the contribution of the present paper enormously and would probably also boost its acceptance and citability in the relevant communities.

Finally, I think the paper would probably benefit most if there were a way to integrate the findings presented here with higher-level theories of cognition, as suggested above. This would shift its status as an almost entirely exploratory paper to a more confirmatory perspective and maybe integrate the paper into one of the current discourses, for example on embodiment, cross-modal cognition or predictive coding as a general cognitive and neural principle (e.g., Clark, 2013).

There is no doubt that the paper carries an important overall take-home message for empirical music researchers: it is possible to make sense of ecologically-valid but very messy data that is generated in response tasks with very few constraints. However, the flip side is that you might need to team up with your local mathematician. Capturing the dynamic and time-series aspects of music perception and cognition is not done often enough in the literature, and this paper introduces a novel and very clever way of describing how our perception and our responses change over time when we listen to music.

REFERENCES

- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, Vol. 36, No. 3, pp. 1-24.
- Schubert, E., & Dunsmuir, W. (1999). Regression modelling continuous data in music psychology. In: S. W. Yi (Ed.), *Music, Mind, and Science*. Seoul: National University Press, pp. 298-352.
- Vines, B.W., Nuzzo, R.L., & Levitin, D.J. (2005). Analyzing temporal dynamics in music: Differential calculus, physics, and functional data analysis techniques. *Music Perception*, Vol. 23, No. 2, pp. 137-152.