Using Linear Mixed-Effects Models to Analyze Historical Trends in Performance Strategies

Commentary on Majid Motavasseli's Article "Interpretation of Cyclic Form in Bach's 'Goldberg Variations' through Performance History"

Bruno Gingras

The usefulness of advanced statistical approaches in performance-based analysis has been demonstrated in several recent publications. Here, I expand on Majid Motavasseli's article "Interpretation of Cyclic Form in J. S. Bach's "Goldberg Variations" through Performance History" (2021) by applying his data to further address issues related to the performance history of the "Goldberg Variations." Using a tempo measurement database comprising seventy-six selected recordings of the "Goldberg" cycle, I built a linear mixed-effects model predicting the main tempo for each recording from the following parameters: performer, variation number, instrument (piano or harpsichord), year of recording, variation type, and finally mode (major/minor) of the variation. Two statistically significant interactions emerged: a two-way interaction between instrument and mode, and a three-way interaction between instrument, year of recording, and variation type. I discuss the musical meaning of these interactions and suggest further avenues for applying advanced statistical approaches in the field of music performance research.

Die Nützlichkeit fortgeschrittener statistischer Ansätze in der Analyse musikalischer Aufführung wurde in mehreren neueren Veröffentlichungen gezeigt. Hier erweitere ich Majid Motavasselis Artikel »Interpretation of Cyclic Form in J. S. Bach's ›Goldberg Variations‹ through Performance History« (2021), indem ich seine Daten anwende, um weitere Fragen im Zusammenhang mit der Aufführungsgeschichte der ›Goldberg-Variationen‹ zu behandeln. Unter Verwendung einer Datenbank mit Tempomessungen von 76 ausgewählten Aufnahmen des ›Goldberg‹-Zyklus habe ich ein lineares Mixed-Effects-Modell erstellt, welches das Haupttempo für jede Aufnahme aus den folgenden Parametern vorhersagt: Interpret*in, Variationsnummer, Instrument (Klavier oder Cembalo), Jahr der Aufnahme, Variationstyp sowie Tongeschlecht der Variation. Es ergaben sich zwei statistisch signifikante Interaktionen: eine Zwei-Wege-Interaktion zwischen Instrument und Tongeschlecht und eine Drei-Wege-Interaktion zwischen Instrument, Aufnahmejahr und Variationstyp. Ich diskutiere die musikalische Bedeutung dieser Interaktionen und schlage weitere Wege zur Anwendung fortgeschrittener statistischer Ansätze im Bereich der Musikperformanceforschung vor.

Schlagworte/Keywords: corpus-based research; cyclic form; Goldberg Variations BWV 988; Goldberg-Variationen BWV 988; Johann Sebastian Bach; korpusbasierte Forschung; performance analysis; Performance-Analyse; statistical-quantitative methods of music analysis; statistisch-quantitative Analysemethoden; zyklische Form

In this commentary, I would like to expand on Majid Motavasseli's article "Interpretation of Cyclic Form in Bach's 'Goldberg Variations' through Performance History" by applying Motavasseli's data to further address issues related to the performance history of the "Goldberg Variations." More specifically, I will use the tempo measurement database built in the context of the PETAL research project[1] to explore the historical changes in performance tempo in seventy-six selected recordings (from sixty-two different performers) of the "Goldberg Variations" spanning a period from 1928 to 2020.[2] The database includes the name of the performer and year of the recording, a "main tempo" for each of the thirty-two pieces comprising the set of variations (with two measurements for the two different sections of Variation 16, yielding a total of thirty-three tempo values per recording),[3] the instrument employed for the recording (piano or harpsichord), the mode (major/minor), and finally the type of variation (character, virtuoso, or canonic).[4]

Unlike Motavasseli, I will not discuss the tempo relations between pieces in the context of their position within the cycle of variations. Rather, I aim to uncover broader trends over time in the interpretative choices made by performers recording the "Goldberg" cycle. Here, I will limit myself to the analysis of tempo choices, which are the main focus of the PETAL database. More specifically, I will explore the interconnections between tempo and choice of instrument, type of variation, as well as mode.

Another aim of the present article is to familiarize the music-theoretical community with more advanced statistical methods that may help uncover statistical trends that remain difficult to detect using only basic tools such as means, standard deviations, and correlations. In particular, the current article will make use of linear mixed-effects models (LMM), a powerful class of statistical models that are increasingly employed in other fields such as psychology or ecology.[5] LMMs allow for the modeling of both *fixed effects* (typically, variables of interest that are selected by the researcher to include all experimental conditions [or parameter levels] in the study) and *random effects* (typically, a classification parameter used to group together measurements obtained on the same individual [or cluster of individuals] which are randomly sampled from a larger population). In the example at hand, *year of recording*, *instrument employed for the recording*, *mode*, and *type of variation* are treated as fixed effects, whereas *performer* and *piece* are identified as random effects.[6] As this example suggests, LMMs are especially interesting for performance analyses aiming to disentangle statistical effects corresponding to performers'

1 *Performing, Experiencing, and Theorizing Augmented Listening*, project funded by the Austrian Science Fund FWF (P 30058-G26, 01/09/2017–31/08/2020) and located at the University of Music and Performing Arts Graz (KUG), https://petal.kug.ac.at.

2 For more information about the database and the methodology used to obtain the tempo measurements, see Motavasseli 2021 (in this issue). A discography of all analyzed recordings can be found in ibid., Appendix, Table 7.

3 Four tempo measurements of individual variations are missing, three for the 1928 recording by Rudolf Serkin and one for the 1952 recording by Ralph Kirkpatrick, yielding a total of 2504 measurements. Some performers included in the database recorded the cycle more than once, hence the sixty-two different performers for seventy-six recordings. One performer (Lang Lang) made two recordings in 2020; the date of the second recording was arbitrarily changed to 2020.5 to differentiate the two recordings in the statistical models.

4 See Utz 2017, 20, and Williams 2004, 42–43.

5 Laird/Ware 1982.

6 Strictly speaking, the selection of these specific recordings of the *Variations* is not entirely random, but they were chosen out of a much larger population of possible performers and recordings and cannot be categorized using a small number of predefined levels, unlike parameters such as mode or type of variation. For more about random effects, see Baayen/Davidson/Bates 2008.

individual styles from piece-specific effects, something which is difficult to do using more traditional modeling approaches such as analysis of variance (ANOVA).[7]

All the analyses presented here are based on the PETAL database as described above. LMM models were built using the lmer function from package lme4[8] in the R programming environment (version 3.6.1),[9] and the package sjPlot was used to create the graph shown in Figure 1.[10] As indicated above, the goal was to model the main tempo (the dependent variable) using both fixed-effects and random-effects predictors. The initial model included the following predictors: *instrument*, *year of recording*, *mode* of the variation, and *type* of the variation as fixed effects, as well as all possible interactions between these parameters, and *performer* and *variation number* (using a number to identify each individual variation) as random effects. The *step* function of the lmerTest package was used to simplify this initial model by automatically selecting only parameters (including higher-level interactions) that significantly improve the fit of the model to the data.[11]

After applying the *step* function, a final model was obtained, which included two significant higher-level interactions, as determined with a likelihood ratio test: a two-way interaction between *instrument* and *mode* ($\chi^2(1) = 5.5985$, $p < 0.05$), and a three-way interaction between *instrument*, *type*, and *year of recording* ($\chi^2(3) = 15.666$, $p < 0.01$). The two-way interaction is relatively easy to interpret: variations in the minor mode recorded on the piano are noticeably slower than when recorded on the harpsichord,[12] whereas the mean tempo is practically the same for major-mode variations whether they are recorded on the harpsichord or on the piano. This is probably because, while minor-mode variations are played at a slower mean tempo regardless of the instrument, choosing an extremely slow tempo on the harpsichord may lead to a "disjointed" impression given the quick decay associated with this instrument, a particularity which is not as pronounced on the piano.

The three-way interaction between *instrument*, *type*, and *year of recording* is slightly more complex to interpret but can still be visualized fairly easily (Fig. 1). Essentially, the mean tempo varies according to the year of the recording, but not in the same manner according to the instrument and the variation type. More specifically, Figure 1 shows that, for both instruments, virtuoso variations (in purple) are played at a slower tempo in more recent recordings. The same trend is visible for the *Aria* (in red), although much more pronounced for piano recordings than for harpsichord recordings (the *Aria* is generally played at a slow tempo and, as mentioned above, extremely slow tempi may not be appropriate for the harpsichord). For the canonic variations (in blue), the tempo also slows down in more recent recordings, but only very slightly (the trend is similar for both piano and harpsichord recordings). Finally, and interestingly, characteristic variations (in green) tend to be played faster in more recent harpsichord recordings, while they are played slower on the piano recordings. To summarize, there is a general tendency to play the

7 See Gingras/Asselin/McAdams 2013.

8 See Bates/Maechler/Bolker/Walker 2015.

9 See R Core Team 2019.

10 See Lüdecke 2018.

11 See Kuznetsova/Brockhoff/Christensen 2017.

12 See Motavasseli 2021, Figure 2, for a comparison of tempi for variation 25 (one of the three minor variations) by harpsichordists and pianists.

"Goldberg" cycle slower in more recent recordings, but this tendency is much more pronounced in piano recordings than in harpsichord ones, and also more pronounced for the virtuoso variations and for the *Aria* than for the other variation types (the exception being the characteristic variations which are actually played faster in recent harpsichord recordings).

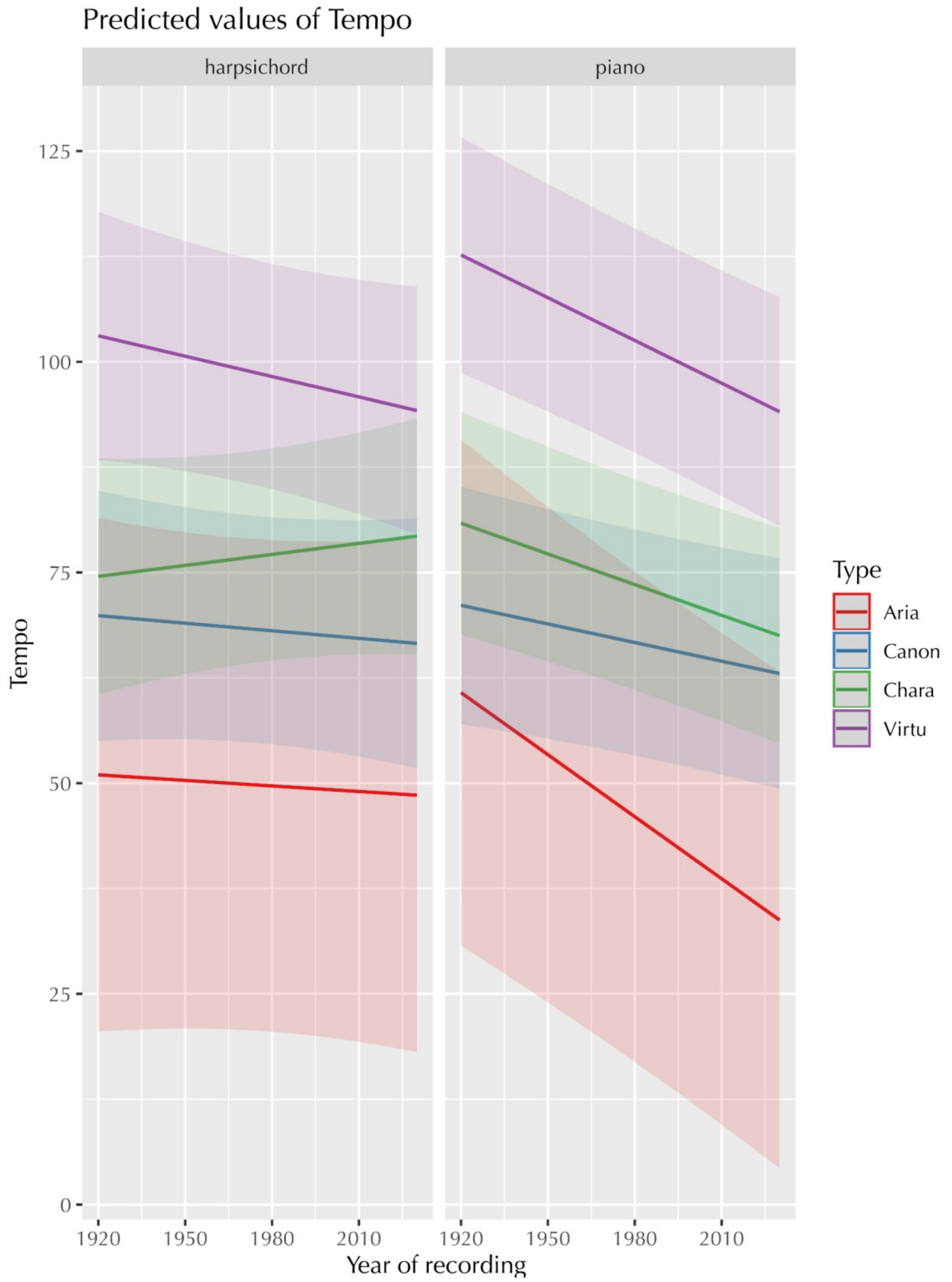

Figure 1: Visualization of the three-way interaction between instrument, year of recording, and variation type. Aria: opening and closing *Arias*; Chara: characteristic variations; Canon: canonic variations; Virtu: virtuoso variations; lighter bands represent the 95% confidence intervals for each estimate.

It is worth noting that the PETAL dataset includes performances with some extreme or otherwise atypical tempo values, such as a few exceptionally fast piano-roll recordings by Rudolf Serkin (1928). To evaluate the potential impact of these outliers, I used the influence.ME package in R, which identified seventy-two influential recordings out of 2504 with a Cook's distance larger than 4/*n* (with *n* being the total number of observations, 2504 in this case).[13] I then repeated the analysis on a subset of the PETAL dataset excluding these seventy-two "extreme" recordings (most of them from Glenn Gould, Ralph Kirkpatrick, Wanda Landowska, Rudolf Serkin, and Yūji Takahashi). I obtained the same LMM model as described above, with only minor differences in the *p*-values and coefficient estimates. This suggests that the statistical trends reported here are not driven by a few atypical recordings but are indeed generally representative of the entire PETAL dataset for the "Goldberg" cycle.

It would be beyond the scope of this commentary to speculate on the reasons and motives behind these changes in performance strategies over time, as well as on the complex interplay between tempo, choice of instrument, and variation type (with an additional interaction between instrument and mode). Nevertheless, the observation that these statistical trends exist might, in itself, be a starting point for further musicological and performance-based analysis along the lines of Motavasseli's article. On a related note, it should be pointed out that the analysis presented here was not based on a dataset that was selected for the express purpose of confirming a preconceived hypothesis. Rather, the findings reported here are entirely the products of an objective, principle-based statistical approach, and emerged serendipitously, as it were, from an analysis of the PETAL database.

As mentioned at the outset of this commentary, the analysis reported here does not address temporal relationships between individual variations, at least not beyond general trends related to variation type and mode. However, although such an analysis falls outside the purview of this essay, it is also possible to apply more sophisticated statistical models that take into account the order of the variations within the cycle as well as other possible interconnections within the various pieces. For instance, autoregressive models, notably time-series analyses, have been fruitfully applied to analyze tempo variations within a single piece, taking into account the order in which the notes are played, as well as the relationship between tempo and other factors.[14] While these models are somewhat more complex than LMM models, they would allow a statistically principled analysis of the temporal relationships explored in Motavasseli's article. It is to be hoped that the brief outline presented here will inspire researchers in the field of performance-based analysis to further familiarize themselves with these statistical approaches that can help reveal rich interconnections between various factors that would otherwise remain concealed.

13 Cook's distance is an estimate of the influence of a data point. For the 4/*n* rule, see Bollen/Jackman 1990. For the influence.ME package, see Nienwenhuis/Grotenhuis/Pelzer 2012.

14 See Gingras et al 2016.

References

Baayen, R. Harald / Douglas J. Davidson / Douglas M. Bates. 2008. "Mixed-Effects Modeling with Crossed Random Effects for Subjects and Items." *Journal of Memory and Language* 59/4: 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bates, Douglas / Martin Mächler / Ben Bolker / Steve Walker. 2015. "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software* 67/1: 1–48. https://doi.org/10.18637/jss.v067.i01

Bollen, Kenneth A. / Robert W. Jackman. 1990. "Regression Diagnostics: An Expository Treatment of Outliers and Influential Cases." In *Modern Methods of Data Analysis*, edited by John Fox and J. Scott Long. Newbury Park, CA: Sage, 257–291.

Gingras, Bruno / Pierre-Yves Asselin / Stephen McAdams. 2013. "Individuality in Harpsichord Performance: Disentangling Performer- and Piece-Specific Influences on Interpretive Choices." *Frontiers in Psychology* 4/895. https://doi.org/10.3389/fpsyg.2013.00895

Gingras, Bruno / Marcus T. Pearce / Meghan Goodchild / Roger T. Dean / Geraint Wiggins / Stephen McAdams. 2016. "Linking Melodic Expectation to Expressive Performance Timing and Perceived Musical Tension." *Journal of Experimental Psychology* 42/4: 594–609. https://doi.org/10.1037/xhp0000141

Kuznetsova, Alexandra / Per B. Brockhoff / Rune H. B. Christensen. 2017. "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software* 82/13: 1–26. https://doi.org/10.18637/jss.v082.i13

Lüdecke, Daniel. 2018. "sjPlot: Data Visualization for Statistics in Social Science. R package version 2.8.7." https://CRAN.R-project.org/package=sjPlot.

Motavasseli, Majid. 2021. "Interpretation of Cyclic Form in J. S. Bach's 'Goldberg Variations' through Performance History." *Zeitschrift der Gesellschaft für Musiktheorie* 18, Special Issue: *Musikalische Interpretation als Analyse. Historische, empirische und analytische Annäherungen an Aufführungsstrategien musikalischer Zyklen*, 19–69. https://doi.org/10.31751/1119

Nieuwenhuis, Rense / Manfred te Grotenhuis / Ben Pelzer. 2012. "influence.ME: Tools for Detecting Influential Data in Mixed Effects Models." *The R Journal* 4/2: 38–47. https://doi.org/10.31235/osf.io/a5w4u

R Core Team. 2019. *The R Project for Statistical Computing*. Vienna: R Foundation for Statistical Computing. http://www.R-project.org

Utz, Christian. 2017. "Komponierte, interpretierte und wahrgenommene Zeit. Zur Integration temporaler Strukturen in eine performative Analyse – eine Diskussion anhand von Johann Sebastian Bachs *Goldberg-Variationen*." *Musik & Ästhetik* 21/82: 5–23.

Williams, Peter. 2004. *Bach: The "Goldberg Variations*." Cambridge: Cambridge University Press.

Universität Innsbruck [University of Innsbruck]

Gingras, Bruno (2021), "Using Linear Mixed-Effects Models to Analyze Historical Trends in Performance Strategies: Commentary on Majid Motavasseli's Article 'Interpretation of Cyclic Form in Bach's 'Goldberg Variations' through Performance History'", *Zeitschrift der Gesellschaft für Musiktheorie* 18/Sonderausgabe [Special Issue]: *Musikalische Interpretation als Analyse. Historische, empirische und analytische Annäherungen an Aufführungsstrategien musikalischer Zyklen*, 71–77. https://doi.org/10.31751/1120

eingereicht / submitted: 15/02/2021

angenommen / accepted: 05/07/2021

veröffentlicht / first published: 05/11/2021

zuletzt geändert / last updated: 08/10/2021