Volume 10, Number 3, September 2004 Copyright © 2004 Society for Music Theory

Rules for Chord Doubling (and Spacing): A Reply to Wibberley*

Paul T. von Hippel and Bret Aarden



REFERENCE: http://mtosmt.org/issues/mto.04.10.3/mto.04.10.3.wibberley.php

KEYWORDS: pedagogy, doubling, spacing, rules, chord

ABSTRACT: In MTO 10.2, we systematically tested common-practice "Rules for Chord Doubling (and Spacing)." We concluded that, while most of the commonly taught rules are not particularly helpful, two are definitely worth teaching: students should be taught to avoid doubling tendency tones, and to favor wide spacing in the bass. In the current issue of MTO, Wibberley expresses an ambivalent attitude toward rules, and questions our methods and results. Wibberley's criticism, however, is based on a sample of just 10 chord pairs, which (as it turns out) are unrepresentative of the 3603 chord pairs in our analysis. We clarify our methods and results, respond to some questions raised by Wibberley and other readers, and discuss the merits and limits of musical rules.

Received August 2004

[1] In the last issue of MTO, (1) we described a method for testing rules of chord doubling and spacing. In this method, a "composed" chord, extracted from the works of Bach, Haydn, and Mozart, is paired with a "random" chord sampled from the set of all possible chords that satisfy certain constraints. (2) Both the "composed" and "random" chords are then inspected for features that are consistent with doubling and spacing rules. If a rule is correct, then rule-consistent features should be more common in the "composed" chords than in the "random" chords.

[2] It may be helpful to imagine that the "random" chords were written by students who have not yet been taught doubling or spacing rules. In this scenario, our task is to differentiate between "student practice" (random chords) and "historical practice" (composed chords). Our goal is to determine which rules will bring these two groups closer together.

- [3] Two well-known rules were verified by this approach. Compared to the random chords,
 - 1. The composed chords were almost four times less likely to double a tendency tone.
 - 2. The composed chords were one-and-a-half times as likely to have their widest space between the two lowest voices.

These results tend to favor the view that Bach, Haydn, and Mozart avoided doubling tendency tones, and favored wide spacing in the bass.

[4] Our conclusion noted that both these preferences are consistent with the goal of creating a transparent musical texture

with distinct melodic lines.⁽³⁾ In particular, the rule against doubling tendency tones can be related to the avoidance of parallel octaves, and the preference for open spacing in the bass is probably related to the auditory system's poorer resolution in the bass region.

- [5] Other rules proved less convincing. Some rules were redundant with the two given above, and some distinguished only weakly between composed and random chords. We were not surprised to see many rules fail, since the doubling literature describes a large number of rules with unclear theoretical motivation.
- [6] On balance, our results suggest that two of the traditional rules are valid and useful summaries of strong common-practice preferences. Other rules are less helpful, and should probably be struck from the curriculum.

Sampling Variation

- [7] Wibberley (this issue) expresses a skeptical view of theoretical rules. In a sample of 10 chord pairs, he reports that a weighted set of rules discriminated composed from random chords only 50% of the time. By contrast, blind guessing identified the composed chord 60% of the time.
- [8] These results are true for the sample of 10 chord pairs that Wibberley examined, but these chords are not very representative. Our results, by contrast, were based on systematic analysis of 3603 chord pairs. In 70% of these pairs, a weighted set of rules discriminated correctly between the composed and random chords. Blind guessing would discriminate only 50% of all pairs.
- [9] The salient issue here is *sampling variation*: if you take a different sample, you get different results. In Wibberley's sample of 10 chords, our model was 50% accurate, and blind guesses were 60% correct. In most samples, however, our model would do better, and guessing would do worse. Readers can check this for themselves by using our paper's interface⁽⁴⁾ to select several samples of chords pair. The model's accuracy will fluctuate from sample to sample, but the fluctuation will be around an average level of 70%. Likewise, blind guessing will fluctuate around its long-term average of 50%. ⁽⁵⁾
- [10] Rather than comparing the model's performance to the performance of one reader on 10 chord pairs, it is more revealing to compare the model's performance to the performance of all readers who have tried to discriminate composed from random chords. As of this writing, 137 readers have tried this discrimination task, and their average accuracy is 57%. Among the 47 readers with PhDs in music, the average accuracy was 58%. In light of these results for human experts, the model's accuracy of 70% seems remarkably good.

Incorrect Rules and Incorrect Answers—but No Incorrect Chords

- [11] Much of Wibberley's comment seems animated by the fear that we are using theoretical rules to evaluate musical practice. If a chord violates a theoretical rule, that is, Wibberley fears that we will label the chord as "incorrect."
- [12] What we are doing, however, is just the opposite: we use musical practice to evaluate theoretical rules. If a composer's chords tend to violate some rule, we consider the rule, not the chords, to be incorrect.
- [13] Since there are no "incorrect" chords, our distinction between "composed" and "random" chords bears repeating. According to the definition in our paper, a "random" chord is not (as Wibberley suggests) an improper or unusable chord. It is rather a chord sampled, at random, from the set of all possible chords that satisfy certain constraints.
- [14] Wibberley points out that several of our "random" chords resemble chords that he can find in chorale harmonizations by Bach. This is true. The fact that a chord was randomly sampled does not prevent a similar chord from being used in a famous composition. Our focus, however, is not on whether a chord occurs, but on how often it occurs. On average, "random" chords occur less often in the repertory than do "composed" chords which were actually sampled from music by Bach, Haydn, and Mozart. And if chords that fit a rule occur more often among the composed chords, then we can infer the rule has merit. This is why rules of musical practice should be helpful for discriminating composed from random chords.
- [15] Because "random" chords can be quite plausible, "composed" and "random" chords cannot be discriminated with 100% accuracy. That is, a chord that we generated randomly will sometimes be taken, quite reasonably, for a chord that we sampled from works by Bach, Haydn, and Mozart. In our paper, we described such failures of discrimination as "incorrect" answers. (6) We emphasize that the word "incorrect" refers only to the discrimination—not to the chords, the model, or the

reader. The reader's understanding may be sound, and one or both chords may be quite typical of the pertinent repertoire—but the discrimination can still fail.

[16] The impossibility of 100% discrimination puts our model's performance in a positive light. Our model discriminates with 70% accuracy, and while this is 30% short of perfect (as Wibberley points out), 70% may be close to the highest level possible. As noted above, 70% is certainly higher than the average accuracy of expert readers.

Rules and Context

[17] Just as some "random" chords can appear typical of musical practice, some "composed" chords can appear rather unusual. Wibberley highlights a chord from a Bach chorale which, taken out of context, seems to have odd spacing. He points out, however, that this chord looks like an excellent choice when returned to its natural surroundings.

[18] We do not dispute this. Our approach is designed to highlight certain similarities among composed chords, whereas Wibberley has chosen to highlight the unique circumstances of a particular chord choice. The difference is merely one of emphasis.

[19] To clarify our position, we agree that each chord chosen by Bach, Haydn, and Mozart serves a particular purpose that is best understood by viewing the chord in context.

[20] We also note, however, that a great many chords serve the *shared* purpose of creating a transparent musical texture with distinct melodic lines. This purpose is well-served by practices codified in certain musical "rules." Among those rules are the two highlighted in our study: (1) the avoidance of doubling tendency tones, and (2) the preference for putting the widest space between the two lowest voices.

[21] Although not inviolate, these rules provide a first approximation to common-practice chord-writing, and a glimpse of insight into the aesthetic goals of common-practice music. Well-tested rules can help students to understand and imitate the music of the past.

Addendum:

[22] Wibberley raises two technical issues that are tangential to the main thread of his paper. We address these points in the following sections:

Sample Size

[23] Wibberley challenges our conclusions on the grounds that we looked at "only" 2643 triads from the Bach chorales. This puzzles us, since 2643 far exceeds the number used in previous studies, and of course we also looked at 960 triads from the string quartets of Haydn and Mozart. By contrast, the norm in textbooks is to demonstrate doubling and spacing rules with a handful of illustrations. Wibberley's own paper focuses on ten chords and four chorale phrases.

[24] Wibberley complains that readers "do not know . . . upon what basis this selection [of chords] was made." Actually, we described the selection criteria in section 4.1 of our paper. We are happy, however, to provide more detail. We began with the Bach chorale section of the MuseData database, which contains "only those [185] chorales for which there is a distinct BWV number (BV 253-438)." (7) After excluding modal chorales and chorales without Melisma harmonic analyses, 126 chorales remained, containing 7616 distinct sonorities, many of which are not functional triads. After restricting the sample to complete strong-beat triads in tonally unambiguous passages, we arrived at the 2643 chorale triads used in our study. In a back-of-the-envelope calculation, Wibberley suggests that there are "at least 11,872 chords" in the Bach chorales, but he begins with the 371 chorales in the Riemenschneider collection, which is not available in computer-readable format, and he does not restrict himself to complete strong-beat triads in tonally unambiguous passages.

Overfitting

Weights and Accuracy

[25] Our study used features of doubling and spacing to discriminate between "composed" and "random" chords. The

results were summarized in two ways:

- 1. Weights. Which features discriminated best? That is, which features should receive the most weight in discrimination?
- 2. Accuracy. When features have been assigned appropriate weights, how accurately do they discriminate between composed and random chords?

[26] In our paper, we used the same set of chords to set weights and to measure discriminatory accuracy. Wibberley objects to this. His concern is legitimate, since using the same data for both purposes can yield inflated estimates of accuracy—a problem known as overfitting. However, overfitting is a substantial problem only when the number of features is comparable to the number of chords. In our study we had thousands of chords and only a few dozen features, so we did not consider the issue worth mentioning.

[27] We did address overfitting in unpublished analyses. Our approach, known as cross-validation, was to split the data into two subsets—a training set that was used to set weights, and a testing set where the weights were used to test discriminatory accuracy. We tried several ways of splitting the data. First, we split it at random. Then we split it by repertoire, so that the chorales were used for training and the quartets were used for testing, or vice versa.

[28] No matter how we split the data, discriminatory accuracy was about 70%—just as we reported in the published paper.

Paul T. von Hippel Ohio State University 300 Bricker Hall, 190 N. Oval Mall Columbus, OH 45210 von-hippel.1@osu.edu

Bret Aarden Ohio State University 300 Bricker Hall, 190 N. Oval Mall Columbus, OH 45210 aarden.1@osu.edu

Footnotes

* We thank David Temperley and David Butler for helpful feedback on an earlier draft.

Return to text

1. Bret Aarden and Paul T. von Hippel, "Rules for Chord Doubling and Spacing: Which Ones Do We Need?" (2004), *Music Theory Online* 10.2.

Return to text

2. Each "random" chord is constrained to have the same pitch-classes and inversion as the "composed" chord with which it is paired. The random and composed chords also have similar range constraints. Further detail on the construction of random chords is provided in Aarden and von Hippel, section 4.2.

Return to text

3. For a discussion of these goals and their relationship to musical rules, see David Huron, "Tone and Voice: A Derivation of the Rules of Voice-leading from Perceptual Principles," *Music Perception*, Vol. 19, No. 1 (2001), pages 1–64, and citations therein.

Return to text

- 4. The interface can be found in Aarden and von Hippel (2004). In the sidebar, click 'Can you beat the computer?' Return to text
- 5. In writing our paper, we tried to anticipate the issue of sampling variation. Our pertinent comments are reproduced at the top of Wibberley's Figure 1, where we write that our model's "performance may be better or worse than average, depending

on the particulars of the sample. Your accuracy, too, may differ from its average level."

Return to text

6. One advantage of online publication is that, even after publication, we can edit language that readers find confusing. In examples of the discrimination task, our original paper described the composed chord as "the correct answer." Since this language was confusing to several readers (including Wibberley), we now describe the same chord simply as "the composed chord." (In Aarden and von Hippel, see the sidebar labeled "Can you beat the computer?")

Return to text

7. Center for Computer-Assisted Research in the Humanities. MuseData editions of Bach's works. http://www.musedata.org/encodings/bach. Palo Alto, CA.

Return to text

Copyright Statement

Copyright © 2004 by the Society for Music Theory. All rights reserved.

[1] Copyrights for individual items published in *Music Theory Online (MTO)* are held by their authors. Items appearing in *MTO* may be saved and stored in electronic or paper form, and may be shared among individuals for purposes of scholarly research or discussion, but may *not* be republished in any form, electronic or print, without prior, written permission from the author(s), and advance notification of the editors of *MTO*.

[2] Any redistributed form of items published in MTO must include the following information in a form appropriate to the medium in which the items are to appear:

This item appeared in *Music Theory Online* in [VOLUME #, ISSUE #] on [DAY/MONTH/YEAR]. It was authored by [FULL NAME, EMAIL ADDRESS], with whose written permission it is reprinted here.

[3] Libraries may archive issues of MTO in electronic or paper form for public access so long as each issue is stored in its entirety, and no access fee is charged. Exceptions to these requirements must be approved in writing by the editors of MTO, who will act in accordance with the decisions of the Society for Music Theory.

This document and all portions thereof are protected by U.S. and international copyright laws. Material contained herein may be copied and/or distributed for research purposes only.

Prepared by Brent Yorgason, Managing Editor and Rebecca Flore and Tahirih Motazedian, Editorial Assistants