
The Effects of Structural Interventions in the First Movement of Mozart's Symphony in G Minor K. 550 on Aesthetic Preference

MITCHELL KARNO & VLADIMIR J. KONEČNI
University of California, San Diego

In a debate in this journal concerning the importance of structure in the effects of musical compositions, Robert Batt suggested the first movement of Mozart's Symphony in G Minor K. 550 as the ideal example on which research could contravene the minimal effects of structural interventions that have been observed in prior research by Konečni. This article reports the results of putting Batt's suggestion to empirical test. As Batt pointed out, the first movement of K. 550 can be broken down into nine segments that can be broadly defined as expository (two themes), developmental, and recapitulative (two themes). The original and four differentially intrusive versions of the piece were presented to 42 college non-music-student subjects and 11 college music-student subjects. On the dimensions of pleasingness, interestingness, and a desire to own a copy of the piece, the original version consistently failed to elicit greater preference than the altered versions did in both subject populations, although all versions were much liked by these nonmusician and musician subjects.

In an article generally concerning the importance of structure for aesthetic appreciation in a number of art forms, Konečni (1984) showed that drastic global structural alterations of Beethoven's string quartets and piano sonatas resulted in no measurable advantage for the original versions. In a later study, Gotlieb and Konečni (1985) examined related issues in what is universally recognized as the pinnacle of structural design achievement of baroque music, namely Johann Sebastian Bach's great Goldberg Variations (Gould, 1955; Terry, 1963). In this research, regardless of instrumentation and playing style (seven different recordings were used), the original was not preferred to the altered versions. The alterations included a randomly scrambled order of the 30 variations between the introductory and concluding arias, a similar ordering except with both arias positioned in the middle of the piece, and the isolation and reordering of two triplet sets of variations. It is important to note that

Requests for reprints may be sent to Vladimir J. Konečni, Department of Psychology (C-009), University of California, San Diego, La Jolla, CA 92093.

although the subjects were not music experts, they certainly had no aversion to baroque music. On the contrary, all the versions received high ratings on the scales of warmth, pleasingness, and interestingness.

In the Winter 1987 issue of *Music Perception*, Batt criticized the tenor of this earlier work. He especially focused on the alleged unsuitableness of the Goldberg Variations as a vehicle to test the importance of structure: "Such works are usually based on a self-contained, and even preexisting, theme; to the extent that each variation has the same form, harmonic and melodic structure, and is in the same key, each variation is also self-sufficient" (Batt, 1987, p. 210). Batt claimed, therefore, that the Goldberg Variations could be altered without negative impact to the natural progression of the work because of the variations' independence.

As a proper test, Batt suggested the first movement of Mozart's Symphony in G Minor K. 550. As he pointed out, "altering the order of sections within a movement will permit the investigation of music in which the complete spectrum of tonal channels of structure is operative" (Batt, 1987, p. 212). Furthermore, "Not only could the order of the exposition, development, and recapitulation be altered, but the order of the four sections within the exposition and, separately, within the recapitulation could be altered" (Batt, 1987, p. 212). Clearly, Batt implied that this work would be ideal for a study of this nature.

In the rebuttal to Batt's article, Konečni (1987) outlined an experiment that would put Batt's suggestion to the test (see footnote 1, p. 215) and invited Batt to do it. Since this has not been forthcoming, we have done it ourselves.

Study 1

METHOD

Stimulus Materials

Five versions (the original and four "doctored" ones) of the first movement of Mozart's Symphony in G Minor (no. 40) K. 550¹ were used in this study (Table 1). In Version 2, the order of the themes within both exposition and recapitulation was changed, and in Version 3 all nine sections were randomly arranged. In Version 4, the order of sections within exposition and recapitulation was random, and in Version 5 the order of the exposition, development, and recapitulation was rearranged while leaving the order of the themes in each intact. In making Versions 2 through 5, every attempt was made to reduce signs of modification.

Rating Instruments

Subjects evaluated each variation of the piece on the following three 200-mm bipolar scales: pleasing/not pleasing, wish to own/do not wish to own a recording, interesting/not

1. The piece was performed by the Orchestra of the 18th Century on period instruments and was conducted by Frans Bruggen. Philips recording, 5/1985.

interesting. The positive and negative extremes of the various scales were alternated on the answer sheets to minimize bias. Subjects made their evaluations upon completion of each variation. These evaluations were then measured to the nearest millimeter for the data analysis. In addition, subjects reported which variation they perceived to have the best structure.

Subjects and Experimental Design

Subjects for this study were a total of 42 male and female undergraduate students from the University of California at San Diego.² Subject comments on rating sheets indicated they had little to no knowledge of the symphony movement. A within-subjects design was used in which each version was heard by every subject. To minimize order (especially primacy and recency) effects, we randomly assigned subjects to one of three different groups (Table 2). The music was played at a comfortable listening level. No title was given to the work.

TABLE 1
Structural Arrangements in the Five Experimental Versions

Version	Elements
Original	Exposition: theme 1, transition, theme 2, closing, Development, Recapitulation: theme 1, transition, theme 2, closing
Version 2	Exp.: theme 2, transition, theme 1, closing, Dev., Recap.: theme 2, transition, theme 1, closing
Version 3	Recap.: theme 1, Exp.: theme 1, closing, Dev., Recap.: transition, closing, Exp.: transition, Recap.: theme 2, Exp.: theme 2
Version 4	Exp.: transition, theme 1, closing, theme 2, Dev., Recap.: theme 1, closing, transition, theme 2
Version 5	Dev., Recap.: theme 1, transition, theme 2, closing, Exp.: theme 1, transition, theme 2, closing

NOTE: The specific bar numbering of the first movement of K. 550, as specified by Batt, is as follows: Exposition: theme 1 (mm. 1–20), transition (mm. 21–43), theme 2 (mm. 44–72), closing (mm. 72–99); Development (mm. 100–163); Recapitulation: theme 1 (mm. 164–183), transition (mm. 184–226), theme 2 (mm. 227–260), closing (mm. 260–299).

TABLE 2
Order of Versions Heard by Each Non–Music-Student
Experimental Group

Order	Presentation Sequence				
A	Version 1	Version 4	Version 2	Version 3	Version 5
B	Version 3	Version 4	Version 5	Version 1	Version 2
C	Version 2	Version 4	Version 3	Version 5	Version 1

2. We thank David Summers for his assistance in conducting this experiment.

Subjects were told the following:

You will be presented with a series of groupings of music segments. While there will be a slight disruption in the flow of what you hear, please keep your attention focused on the music. After hearing each group of music segments, you will be asked to rate them on four scales [hand out rating sheets.] Please avoid the scale ends on each measure, because you may prevent yourself from responding higher or lower on a subsequent group. Each group of music segments will last approximately 6 min.

RESULTS

Although this composition was singled out by Batt as one the enjoyment of which would be strongly affected by structural modifications, subjects' ratings were minimally influenced by the drastic alterations. The results are presented in Tables 3–5. There were no significant main effects for either the Order or Version factors. However, there were highly significant interactions between these two factors on all three measures (see notes to Tables 3–5). These interactions can be clearly interpreted as strong primacy effects. When the original version was heard first (Order A), subjects rated it higher than all four altered versions on each of the three dependent measures. Furthermore, the original was significantly preferred over only three of the four other versions, specifically, those heard second, fourth, and fifth (using Duncan's Multiple Range test, Original > Version 4, $df(4,65)$, $p < .05$; Original > Version 3, $df(3,65)$, $p < .05$; and Original > Version 5, $df(5,65)$, $p < .05$, respectively).

However, these primacy effects were just as strong for versions other than the original. Subjects listening to Orders B and C—where the version heard first was not the original—also rated the version heard first the highest (Versions 3 and 2 were heard first in Orders B and C, respectively).

TABLE 3
Treatment Means (Standard Deviations) for "Pleasingsness"
Measure of Non-Music-Student Subjects

Order	Version					Mean Totals
	1	2	3	4	5	
A	134.69 (46.13)	114.00 (48.80)	111.92 (50.05)	105.08 (57.74)	91.46 (57.69)	111.43 (52.58)
B	100.57 (52.52)	96.50 (53.05)	121.57 (28.53)	118.07 (30.92)	105.79 (39.23)	108.50 (42.07)
C	120.27 (43.36)	141.00 (32.89)	109.60 (49.14)	112.47 (38.18)	126.87 (48.69)	122.04 (43.24)
Mean Totals	118.17 (48.31)	117.81 (48.12)	114.31 (42.94)	112.05 (42.49)	108.88 (49.89)	114.24

NOTE: The higher the number (0–200), the more pleasing the version was found by the subjects. Interaction: $F(8, 156) = 4.095$, $p = .0002$.

TABLE 4
Treatment Means (Standard Deviations) for "Interestingness" Measure
of Non-Music-Student Subjects

Order	Version					Mean Totals
	1	2	3	4	5	
A	138.69 (39.28)	112.31 (51.27)	107.62 (51.39)	101.23 (53.30)	92.08 (56.94)	110.39 (51.65)
B	95.36 (49.05)	90.29 (47.54)	127.71 (32.00)	123.21 (30.40)	114.14 (27.19)	110.14 (40.13)
C	109.13 (47.62)	126.07 (40.01)	94.13 (49.15)	117.47 (35.66)	105.93 (49.11)	110.55 (44.76)
Mean Totals	113.69 (48.06)	109.88 (47.55)	109.50 (46.08)	114.36 (40.54)	104.38 (45.70)	110.36

NOTE. The higher the number (0–200), the more the subjects found the version interesting. Interaction: $F(8, 156) = 5.464, p = .0001$.

TABLE 5
Treatment Means (Standard Deviations) for "Desire to Own" Measure
of Non-Music-Student Subjects

Order	Version					Mean Totals
	1	2	3	4	5	
A	104.77 (55.79)	86.54 (50.16)	80.46 (50.62)	81.23 (57.66)	70.85 (53.70)	84.77 (53.17)
B	77.21 (43.70)	73.86 (44.70)	90.43 (39.11)	87.79 (38.76)	85.57 (37.17)	82.97 (40.11)
C	86.13 (53.51)	106.27 (45.51)	78.67 (57.35)	76.33 (48.51)	92.07 (56.86)	87.89 (52.25)
Mean Totals	88.93 (51.22)	89.36 (47.60)	83.36 (48.84)	81.67 (47.71)	83.33 (49.65)	85.29

NOTE. The higher the number (0–200), the greater the subjects' desire to own a recording. Interaction: $F(8, 156) = 4.026, p = .0002$.

It is important to keep in mind that Version 3 was a random arrangement of all nine sections; thus, it presented the greatest opportunity for the detection of structural changes. Still, it was rated highest overall in Order B and was significantly preferred over another nonoriginal, yet much less altered version (by Duncan's test, Version 3 > Version 2, $df(5,65), p < .05$).

The pattern of the data refutes Batt's claims by demonstrating that preference for a particular version relies little on structure and heavily on primacy.

TABLE 6
**Breakdown of Non-Music-Student Subjects' Ratings
 on "Best Overall Structure" Measure**

Order	Version				
	1	2	3	4	5
A	30.77%	23.08%	30.77%	7.69%	7.69%
B	35.71%	0.00%	28.57%	14.29%	21.43%
C	20.00%	20.00%	0.00%	13.33%	46.67%
Mean Totals	28.57%	14.29%	19.05%	11.90%	26.19%

NOTE. The higher the percentage, the greater the number of respondents who favored that particular version.

Results from the fourth dependent measure, "best overall structure," are presented in Table 6. Although the original was preferred over the other versions, chi-square analysis revealed these differences were insignificant, $\chi^2(8) = 12.67, p = .124$. Interestingly, Version 2, the least altered of the four nonoriginal versions, ranked next to lowest, while Version 5, despite rather extensive structural manipulations, ranked just below the original. These findings imply that the original was by no means the undisputed favorite and that factors unrelated to structure were influencing the subjects' responses.

Study 2

A second study was conducted to examine further the importance of structure in a musical work. By using primarily the same methods as in Study 1, this study followed Batt's suggestion to test those individuals educated in music. Since music students are arguably more sensitive to structural alterations in music than are non-music students, their employment as subjects gives structure the greatest opportunity to demonstrate its importance.

METHOD

Stimulus Materials

Recordings used for Study 2 were identical to those used in Study 1 (See Table 1).

Rating Instrument

Measurements used for Study 2 were identical to those used in Study 1.

Subjects and Experimental Design

Subjects for this study were 11 undergraduate music majors from the University of California at San Diego. Subject comments on rating sheets indicated that recognition for the symphony movement was slightly higher than that seen in Study 1; one music-student subject knew the movement well. A within-subjects design was used in which each version was heard by every subject. To minimize order (especially primacy and recency) effects, we randomly assigned subjects to one of two groups (Table 7). The music was played at a comfortable listening level. No title was given to the work. Subject instructions were the same as used in Study 1.

RESULTS

Despite claims by Batt about the increased sensitivity of musically educated individuals to structural changes in musical works, subjects' ratings failed to reflect any heightened awareness. The results are presented in Tables 8–10. As in Study 1, there were no significant main effects for either the Order or Version factors. Moreover, no significant interactions were present on any of the three measures. These findings clearly refute Batt's claims by further demonstrating that preference for a particular version does not rely on structure.

TABLE 7
Order of Versions Heard by Each Music-Student
Experimental Group

Order	Presentation Sequence				
A	Version 1	Version 4	Version 2	Version 3	Version 5
B	Version 3	Version 4	Version 5	Version 1	Version 2

TABLE 8
Treatment Means (Standard Deviations) for "Pleasantness" Measure of
Music-Student Subjects

Order	Version					Mean Totals
	1	2	3	4	5	
A	156.33 (27.92)	133.50 (44.75)	132.50 (58.52)	147.67 (26.55)	122.67 (69.87)	138.53 (46.70)
B	142.40 (24.06)	127.20 (22.78)	142.40 (24.06)	132.20 (34.49)	141.00 (18.01)	137.04 (23.92)
Mean Totals	150.00 (25.97)	130.64 (34.92)	137.00 (44.39)	140.64 (29.89)	131.00 (51.60)	137.86

NOTE. The higher the number (0–200), the more pleasing the version was found by the subjects.

TABLE 9
Treatment Means (Standard Deviations) for "Interestingness" Measure
of Music-Student Subjects

Order	Version					Mean Totals
	1	2	3	4	5	
A	162.17 (29.27)	128.67 (46.37)	131.83 (59.98)	139.33 (37.07)	123.83 (70.36)	137.17 (49.17)
B	81.20 (38.86)	77.80 (45.89)	125.80 (22.86)	107.40 (28.96)	100.80 (36.81)	98.60 (37.15)
Mean Totals	125.36 (53.11)	105.55 (51.22)	129.09 (44.92)	124.82 (36.06)	113.36 (56.23)	119.64

NOTE. The higher the number (0–200), the more the subjects found the version interesting.

TABLE 10
Treatment Means (Standard Deviations) for "Desire to Own" Measure
of Music-Student Subjects

Order	Version					Mean Totals
	1	2	3	4	5	
A	117.33 (66.03)	100.00 (56.52)	103.17 (62.77)	109.50 (53.34)	93.17 (70.66)	104.63 (58.34)
B	79.80 (51.77)	69.40 (46.19)	119.80 (35.48)	105.20 (39.77)	99.00 (32.12)	94.64 (43.40)
Mean Totals	100.27 (62.22)	86.09 (52.02)	110.73 (50.49)	107.55 (45.39)	95.82 (54.02)	100.09

NOTE. The higher the number (0–200), the greater the subjects' desire to own a recording.

TABLE 11
Breakdown of Music-Student Subjects' Ratings on "Best Overall
Structure" Measure

Order	Version				
	1	2	3	4	5
A	33.33%	0.00%	16.67%	16.67%	33.33%
B	0.00%	40.00%	40.00%	0.00%	20.00%
Mean Totals	18.18%	18.18%	27.27%	9.09%	27.27%

NOTE. The higher the percentage, the greater the number of respondents who favored that particular version.

Results from the “best overall structure” measure are presented in Table 11. The original was clearly not preferred over all other versions. Furthermore, chi-square analysis found no significant differences between subject ratings, $\chi^2(4) = 5.62$, $p = .23$. Surprisingly, Versions 5 and 3 (a completely random structural arrangement) ranked highest overall. Again it appears that even among music students the original was not the most favored and that factors other than structure were affecting subjects’ preferences.

DISCUSSION

Contrary to the claims of music scholars and music historians, the ability of even musically trained individuals to perceive structural changes is highly suspect. Study 2 has demonstrated that subjects possessing theoretical and practical knowledge of structure in music are no more aware of it as an audience than are subjects with no formal music training.

Admittedly, 11 subjects is not necessarily an effective representation of the music-student population. One must remember, however, that Batt implied that any structural changes in the symphony movement used in this study would result in a *significantly inferior piece*. Were these altered versions truly inferior, differences would appear in ratings between the original and nonoriginal versions among even a few musically trained subjects.

General Discussion

Despite its being directly compared with versions having undergone drastic structural alterations, the original version of K. 550 consistently failed to elicit greater preference on the measures of pleasingness, interestingness, and desire to own a recording among both non-music-student and music-student populations. Only when the original was heard first did it receive the highest rating. Even then it fell short of being significantly preferred to all of the other versions. Furthermore, the primacy effect was operative for altered versions to an equally pronounced degree.

Studies by Cook (1987) on the effects of large-scale tonal closure on listener response have generated somewhat analogous results. Tonal closure is assumed by music theorists to be a critical structural component in a musical work. However, Cook found that even among musically trained subjects, tonal closure was influential only for pieces of very short duration (< 1 min). Therefore, Cook suggested “that the tonal unity of a sonata is of a conceptual rather than perceptual nature, in contrast to the directly perceptible unity of a single phrase” (Cook, 1987, p. 204).

Batt has criticized the use of undergraduate college students as subjects, arguing that they are “the least likely people that would be consumers of art music, that is, the least likely to attend concerts and buy recordings of art music” (Batt, 1987, p. 207). This, in the big picture of classical music loving and listening, is too hasty a dismissal of this subject population. The uniformly high ratings of each version indicate that the subjects, non-music students and music students alike, showed a mature appreciation for the music and were quite suitable as subjects.

It would seem that many (armchair-based) claims made by musicologists about the role of structure in a work’s impact can be called into question by empirical research. First, the effects of structure may be severely constrained by relatively nonmusical elements such as primacy effects and effects of the length of the piece. Second, we agree with Cook’s contention that structure may be less usefully viewed as a perceptible element in music appreciation and more as a conceptual tool of music composition. Such impact that structure does have in music appreciation may be due to its involvement in higher order interactions with other musical variables. The precise form of such interactions is at present unknown and is very difficult to investigate empirically.

References

- Batt, R. Comments on “The effects of instrumentation, playing style, and structure in the Goldberg Variations by Johann Sebastian Bach.” *Music Perception*, 1987, 5(2), 207–213.
- Cook, N. The perception of large-scale tonal closure. *Music Perception*, 1987, 5(2), 197–206.
- Gotlieb, H., & Konečni, V. J. The effects of instrumentation, playing style, and structure in the Goldberg Variations by Johann Sebastian Bach. *Music Perception*, 1985, 3, 87–102.
- Gould, G. Jacket notes. Goldberg Variations by Johann Sebastian Bach, with Glenn Gould. Columbia Records, ML 5060, 1955.
- Konečni, V. J. Elusive effects of artists’ “messages.” In W. R. Crozier & A. J. Chapman (Eds.), *Cognitive processes in the perception of art*. Amsterdam: North-Holland Press, 1984.
- Konečni, V. J. Response to Robert Batt. *Music Perception*, 1987, 5(2), 215–217.
- Terry, C. S. *The Music of Bach*. New York: Dover, 1963.