



HAL
open science

Bach, Mozart, and Beethoven: Sorting piano excerpts based on perceived similarity using DiSTATIS

Rachna Raman, Michael Kriegsman, Hervé Ha Abdi, Barbara Tillmann, W. Jay Dowling

► **To cite this version:**

Rachna Raman, Michael Kriegsman, Hervé Ha Abdi, Barbara Tillmann, W. Jay Dowling. Bach, Mozart, and Beethoven: Sorting piano excerpts based on perceived similarity using DiSTATIS. *New Ideas in Psychology*, 2020, 57, pp.100757. 10.1016/j.newideapsych.2019.100757. hal-02378151

HAL Id: hal-02378151

<https://hal.science/hal-02378151>

Submitted on 10 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Running head: SORTING PIANO EXCERPTS USING DiSTATIS

Bach, Mozart, and Beethoven:

Sorting Piano Excerpts based on Perceived Similarity Using DiSTATIS

Rachna Raman¹, Michael A. Kriegsman¹, Hervé Abdi¹, Barbara Tillmann², & W. Jay Dowling¹

¹The University of Texas at Dallas

²Lyon Neuroscience Research Center, Lyon, France

Correspondence concerning this article should be addressed to:

Dr. Rachna Raman

The University of Texas at Dallas,

800 W. Campbell Road,

Richardson, Texas 75080-3021.

rachnaraman@hotmail.com

1 **Abstract**

2 Sorting tasks can reveal the underlying intuitive structure of a collection of items, in this case
3 musical excerpts. Sorting tasks can be used to compare experts and non-experts without relying
4 on specialized vocabulary, and they tend not to fatigue participants. Here, we used the sorting
5 technique with excerpts from the piano music of Bach, Mozart, and Beethoven. Experiment 1
6 involved sorting 21 MIDI-generated stimuli. Experiment 2 utilized 36 excerpts from recorded
7 performances of four pianists (Arrau, Barenboim, Pirès, and Richter). Each experiment involved
8 2 parts: In Part 1, participants sorted excerpts freely into any number of clusters. In Part 2,
9 participants sorted excerpts into 3 clusters according to whether a single composer could have
10 written the pieces in the group. We divided participants into 3 groups based on music training.
11 We investigated the effects of composer, pianist (in Experiment 2), and music training on
12 sorting. To analyze the data, we applied DiSTATIS, a recent adaptation of multi-dimensional
13 scaling specifically adapted to reveal the perceived dissimilarity among items, as well as to
14 investigate group differences. The results showed an effect of composer in both experiments: In
15 Experiment 1, participants were able to strongly differentiate Beethoven's excerpts from those of
16 the other 2 composers. In Experiment 2, they differentiated Mozart's excerpts from Beethoven's,
17 with Bach falling in between those two. In Experiment 2, participants' sorting decisions were
18 strongly influenced by pianists. Richter's performances of the 3 composers were clustered
19 relatively close to the Mozart region of the solution, indicating their clarity and balance; in
20 contrast, those of Barenboim were clustered in the Beethoven region, indicating their
21 sumptuousness and passion. Training effects were not strong and the highest and lowest
22 expertise groups were differentiated only in the free sorting task of Experiment 2.
23 **Key Words:** constrained sorting, free sorting, musical style, musical training, style perception.

1 **Musical Style and its Perception**

2 Research on the cognitive processing of musical style is now common in the field of
3 music perception and cognition (e.g., Atalay & Placek, 1997; Crump, 2002; Dalla Bella &
4 Peretz, 2005; Eastlund, 1992; Eerola, Järvinen, Louhivuori, & Toiviainen, 2001; Gardner, 1973;
5 Gingras, Lagrandeur-Ponce, Giordano, & McAdams, 2011; Gromko, 1993; Hargreaves & North,
6 1999; Miller, 1979; Storino, Dalmonte, & Baroni, 2007; Tekman & Hortaçsu, 2002; Thorisson,
7 1998; Tyler, 1946; Wedin, 1969; Zivic, Shifres, & Cecchi, 2013). However, earlier research has
8 not yet provided satisfactory answers about the processes underlying style perception. For
9 instance, what features of musical style do listeners perceive in order to categorize the excerpts?
10 Musical style is a complex concept for which a wide range of descriptions has been proposed.
11 Musical style has been defined as a “distinguishing and ordering concept” that “groups examples
12 of music according to similarities between them” (Pascall, 1980). Cope (1991) later defined
13 musical style as “the identifiable characteristics of a composer’s music which are recognizably
14 similar from one work to another.” Dalla Bella and Peretz (2005) more recently described
15 musical style as that which could refer to a particular musical language (e.g., tonal vs. atonal), to
16 the music composed within a particular historical era (e.g., baroque vs. classical), or to a
17 particular composer’s technique (e.g., Mozart’s vs. Beethoven’s). The defining characteristics of
18 style include recurring phrases, specific forms, melodic, harmonic, or rhythmic features, timbre,
19 typical textures, and formal organization (Meyer, 1973; Vignal, 1987). For instance, Gundlach’s
20 (1932) study on the analysis of native American music emphasized the importance of rhythm in
21 categorizing music based on style.

22 Gardner (1973) defined style recognition as “a complex cognitive process which
23 demands monitoring of numerous aspects of a stimulus, avoidance of over-emphasis upon a

24 single facet, and attention to the general expressive features of the work.” Musical style
25 perception can be seen as one example of sophisticated implicit learning processes that lead to
26 nonverbal knowledge just by mere exposure to individual items. Comparable to implicit learning
27 of language or tonal music, listeners seem to become sensitive to structural regularities
28 underlying style (Agres, Abdallah, & Pearce, 2018; Raman & Dowling, 2016; Raman &
29 Dowling, 2017). Although listeners clearly differ in expertise in doing musical tasks related to
30 musical style, such as identifying musical style based on composer or performer, they also share
31 a lot of commonalities in music perception (e.g., Bigand & Poulin-Charronnat, 2006; Gingras et
32 al., 2011; Raman & Dowling, 2017). Daily activities further suggest that listeners, both trained
33 and untrained, are highly sophisticated in recognizing musical styles. For instance, when turning
34 on the radio and listening to music, listeners show remarkable consistency in guessing its
35 musical style (e.g., classical, country, or rock music) or even evaluate finer stylistic differences
36 (e.g., baroque, classical, or romantic periods). And listeners can even easily describe musical
37 excerpts of various styles using similar adjectives (e.g., Hevner, 1936; Miller, 1979; Tekman &
38 Hortaçsu, 2002; Watt & Ash, 1998) and open-ended written descriptions (e.g., Dibben, 2001;
39 Morrison & Yeh, 1999; Thorisson, 1998). Despite the apparent ease of perceptual classification,
40 determining how listeners make such nuanced judgements of style is elusive. As Crump (2002)
41 surmised, an important issue here is to determine the low-level perceptual and high-level abstract
42 information (i.e., cues) that listeners perceive and respond to so as to make judgements of
43 musical style. As early as Hevner (1936), studies have shown that some of the primary features
44 that listeners seem to focus on in understanding a musical piece include mode (major vs. minor),
45 harmony, and rhythm. Musical style, and its perception, thus creates a puzzle: Listeners easily
46 and rapidly recognize the style of an historical period or of a composer, but researchers have

47 been largely unsuccessful at providing theoretical descriptions that capture the characteristics of
48 musical style and its perception.

49 Musical style perceptions are also influenced by listeners' music training and general
50 music listening experiences. Meyer (1989) stated that perception of style seems to be a learned
51 behavior for musically trained listeners, which is the case since understanding the chronological
52 development of various musical styles is a necessary part of their music education. For instance,
53 Miller (1979) found that untrained participants tended to focus more on affective qualities of the
54 music (e.g., playful, carefree, pleasing, etc.), whereas expert participants focused more on
55 musical structure and other technical aspects (e.g., form, dynamics, harmony, etc.). Miller also
56 identified an historical dimension but only for the experts, who were able to categorize styles
57 mostly through differences in composers' use of harmony. On the contrary, several other studies
58 have shown that music novices (e.g., Thorisson, 1998; Tyler, 1946; Wedin, 1969) and
59 nonmusicians (e.g., Eastlund, 1992; Gingras et al., 2011; Gromko, 1993), and even non-Western
60 nonmusicians are sensitive to style recognition (e.g., Dalla Bella & Peretz, 2005), and that
61 perhaps such sensitivity is developed merely due to prolonged exposure to any type of music and
62 the perception of low-level musical cues.

63 In understanding listeners' style perceptions, one particular area of interest is how
64 listeners categorize musical pieces based on stylistic aspects into various genres, such as folk,
65 classical, popular, and so forth. Most of the studies in the field have investigated style perception
66 either via ethnomusicology, wherein researchers quantitatively analyze the prototypical melodies
67 of a particular style of music by studying the statistical distribution of different intervals, pitches,
68 or temporal patterns (e.g., Gundlach, 1932; Zivic et al., 2013), or via the use of trained computer
69 networks and simulations to imitate human performance (e.g., Järvinen, Toiviainen, &

70 Louhivuori, 1999; Smaill & Westhead, 1993). Some researchers have even compared computer
71 simulations versus human performance on style identification and categorization (e.g., Atalay &
72 Placek, 1997; Crump, 2002; Tillmann, Abdi, & Dowling, 2004). Through such analyses,
73 researchers have systematically examined both low-level perceptual characteristics, such as
74 statistical patterns of notes, and high-level abstract features, such as harmony, rhythm, and
75 melody, of music that contribute towards style perception.

76 Studies using sorting tasks, which are implicit and do not require verbalization of stylistic
77 features by human listeners, are relatively rare, mostly due to the difficulty in finding suitable
78 methods to accurately measure listeners' perception of style. Earlier studies mostly relied on
79 tasks where participants had to explicitly verbalize their judgements of style (e.g., Tyler, 1946;
80 Gardner, 1973). One of the earliest studies on style perception was conducted by Tyler (1946),
81 who had novice music students listen to 3-min excerpts of three selections each from Mozart,
82 Beethoven, and Schubert. These excerpts were from different movements within the same piece.
83 The stimuli were presented randomly and the students had to verbally judge who the composer
84 of each selection was. The study was performed twice by the same participants on two different
85 occasions during the semester. The results indicated that on both occasions participants showed
86 sensitivity to style recognition, and this sensitivity was related to their prior music training and
87 concert experience, but not with their intelligence scores or their preference of the composers or
88 music pieces.

89 Gardner (1973) conducted a developmental investigation of style, and the obtained data
90 from a verbalization task showcased the importance of participant age in musical style
91 perception. Five groups of participants (ages 6, 8, 11, 14, and 18 – 19 years) judged whether
92 pairs of 15-s excerpts of classical music from 1680 to 1960 belonged to the same piece or not.

93 Participants 11 years and older performed similarly and more accurately than the younger
94 participants, who also showed some sensitivity to such discriminations. Also, participants 11
95 years and older could discriminate pieces based on subtler aspects of style, and were able to
96 categorize the pieces as either from the same musical era or from different eras.

97 With advancement in statistical methods, researchers began using nonverbal “rating”
98 tasks, such as the similarity-judgement task, and multidimensional scaling (MDS) to analyze
99 subjective ratings, thus popularizing its application to investigate recognition of musical style.
100 The similarity-judgement task involves nonverbal categorization of musical stimuli via rating
101 responses. For instance, Wedin (1969) investigated the perceptual dimensions into which the
102 historical epochs of music would fall using MDS. Participants heard 98 pairs of 10-s excerpts
103 (including repetitions) of Western classical music from 1720 to 1890, and rated them in two
104 types of similarity-judgement tasks. In the first task, novices with less than 5 years of music
105 training judged the degree of similarity in percentage between pairs of excerpts. In the second
106 task, a similar procedure was followed and three participant groups (highly and moderately
107 trained groups, and novices) had to rate on a 10-point scale the subjective similarity in style of
108 the pair of excerpts. The results revealed that all participants showed style sensitivity but
109 participants with greater musical training grouped the musical excerpts into four distinct
110 clusters—Baroque, Rococo, Viennese Classicism, and Romanticism. Thus, they showed a
111 clearer and nuanced distinction among the categories. On the contrary, participants with lesser or
112 no musical training grouped the musical excerpts into three distinct clusters—Baroque, Viennese
113 Classicism, and Romanticism.

114 Eastlund (1992; also Gromko, 1993) further extended this approach using the same
115 nonverbal task to investigate differences in style perception among untrained participants, music

116 undergraduates, and experts (music professors). She used 15 musical pieces belonging to either
117 the classical or the romantic styles, composed between 1762 and 1896. Participants heard 105
118 pairings of 15-s music excerpts and performed a similarity-judgement task using a 7-point scale.
119 MDS analysis showed that music undergraduates and expert musicians performed almost
120 identically, and their combined responses could be classified into three dimensions (in order of
121 importance): historical period in which the piece was composed, perceived complexity of the
122 excerpt, and its tempo. On the contrary, for untrained participants historical period was the least
123 important, which partially explains Miller's (1979) findings of a lack of historical dimension for
124 untrained participants. Eastlund proposed an explanation for this difference in style perception,
125 that untrained participants may focus more on what she called secondary features of music (e.g.,
126 tempo, pitch, dynamics, etc.), whereas trained participants focus more on primary features (e.g.,
127 melody, harmony, etc.).

128 Later, Thorisson (1998) examined the validity of style-categorization results from the
129 nonverbal similarity-judgement task by comparing with participants' open-ended written
130 descriptions. He examined whether novice listeners were able to classify musical excerpts as
131 either classical or romantic, based on compositional styles. Participants were first familiarized
132 with 17 classical and romantic piano excerpts, and then they completed similarity ratings of the
133 136 possible pairings of the excerpts. MDS indicated that the excerpts were generally grouped
134 into two clusters, one for the classical period and the other for the romantic era. Listeners gave
135 written descriptions of attributes pertaining to texture, tempo, dynamics, and so forth, for each
136 piece, and the results showed that excerpts from the same musical period but by different
137 composers received similar attributes, thus validating the use of both tasks.

138 Several studies have used the nonverbal similarity-judgement task to determine the exact
139 nature of the musical cues listeners employ to identify the genre of a musical piece. For example,
140 Eerola, Järvinen, Louhivuori, and Toiviainen (2001) used similarity ratings of folk melodies to
141 predict music students' classification of melodies that represented five different European folk
142 styles. MDS analysis showed that the students were able to categorize the melodies based on the
143 different folk styles. The results also indicated that the salient aspects of a musical piece, such as
144 statistical properties of the pitches and rhythm, to which listeners generally paid attention,
145 presumably helped listeners classify the pieces according to melodic similarity.

146 Supplementing Eerola et al.'s (2001) findings, Tekman and Hortaçsu (2002) used a
147 verbalization task to determine how listeners perceived the relationship among various musical
148 styles. They asked Turkish undergraduate students to list all the genres of music they knew and
149 had them rate a list of adjectives on how appropriately they described each genre. MDS
150 identified two components based on the students' classification of the different styles: historical
151 novelty (traditional vs. modern) and appeal (to large population vs. to small groups) of the styles.
152 The students classified closely associated styles, such as rap and techno, systematically.
153 Qualitative analyses showed that the students also described the musical styles on the basis of
154 three dimensions—evaluative, activity, and peacefulness. Tekman and Hortaçsu's study provided
155 evidence that listeners not only possess knowledge on various styles of music but also on the
156 relationship of the styles to each other, and the unique descriptive qualities associated with each
157 style.

158 Later, Dalla Bella and Peretz (2005) investigated the recognition of musical style using a
159 different approach with the same nonverbal similarity-judgement task. They had two
160 professional composers create 16 piano excerpts imitating the musical styles of baroque,

161 classical, romantic, and post-romantic eras, and six advanced piano students recorded the
162 excerpts. Each student played excerpts from only one style so that confusion between
163 compositional styles and performance styles could be avoided. Since the excerpts were
164 composed specifically for the study, the researchers could control for confounds (e.g., familiarity
165 with the musical piece). Western music students, and Western and non-Western (Chinese) non-
166 music students performed a familiarity-rating task for each excerpt and a similarity-judgement
167 task for 128 excerpt pairs. Half of the excerpt pairs were presented in the historical order (e.g.,
168 classical followed by post-romantic) whereas the other half of the pairs were presented in the
169 inverse order (e.g., post-romantic followed by classical). MDS analysis showed that all
170 participants rated melodies from earlier historical periods (e.g., baroque) as more familiar, and
171 they rated compositions closer in styles (i.e., historical eras) as similar. This sensitivity to style
172 recognition was enhanced in the musician group, though both Western and non-Western
173 nonmusician groups also showed an obvious sensitivity, indicating the significance of mere
174 passive long-term exposure to music. Dalla Bella and Peretz also proposed that universal low-
175 level perceptual processes (such as, temporal regularities) may underlie style sensitivity. And
176 finally, the results showed an order effect, wherein participants differentiated the styles more
177 easily when the excerpts were presented in chronological order rather than when reversed.

178 Storino, Dalmonte, and Baroni (2007) further extended Dalla Bella and Peretz's (2005)
179 research by investigating whether familiarization with a single composer's musical grammar can
180 facilitate listeners' style categorization based on that composer's technique and of the
181 corresponding historical period in general. In Experiment 1, expert musicians in the baroque
182 style were first familiarized with eight Legrenzi's (an Italian baroque composer) arias. In the test
183 phase, participants heard excerpts from 10 arias by Legrenzi and 10 arias produced by LEGRE (a

184 computer program that produces new arias in Legrenzi's style for the same texts of music). Half
185 of the participants heard the excerpts as 10 Legrenzi-LEGRE pairs and the remaining
186 participants heard the 20 excerpts in isolation (i.e., not paired). Both groups had to identify
187 which of the excerpts was composed by Legrenzi. All participants were able to classify based on
188 style, however, accuracy was higher in the paired condition and was just above chance level in
189 the isolated condition. In Experiment 2, trained (not in any particular style) and untrained
190 participants performed only the paired task. The results showed that only the trained participants
191 were able to perform above chance level. The method of Experiment 4 was the inverse of
192 Experiment 1, wherein trained participants were first familiarized with eight arias produced by
193 LEGRE. In the test phase, they heard only 18 arias by three Italian baroque composers—
194 Legrenzi, Rossi, and Gabrielli—and not those produced by LEGRE. Participants had to indicate
195 whether the arias had been created by the same composer in the familiarization phase (i.e.,
196 LEGRE) or not, and the results showed that participants were successful, thus confirming the
197 similarity between LEGRE and Legrenzi styles of composition. Storino et al. found that with
198 brief exposure, even musicians non-experts in the baroque style (as in Experiments 2 & 4) were
199 able to perceive stylistic features of the music.

200 Storino et al. (2007) used a sophisticated grammar based on musicological analysis in
201 their LEGRE program. A contrasting approach in recent years relies on simply capturing the
202 sequential regularities of the musical style in the music in a Markov-based statistical learning
203 program (see Agres et al., 2018, for a review). Listeners have been shown to extract the
204 sequential regularities of melodies they were exposed to, and expectancies are fairly well
205 matched by the statistical learning programs. However, as Krumhansl (2015) points out, citing
206 Meyer (1989), there is more to musical style and listeners'—even untrained listeners'—

207 understanding of it than can be captured with Markovian statistical learning. Music has structural
208 and expressive properties that are easily grasped by the attentive listener, but which are not taken
209 into account by Markovian statistics.

210 To summarize, researchers over the years have used various types of verbal and
211 nonverbal tasks in order to ascertain how listeners perceive and categorize stylistic aspects of
212 music. The results show a broad picture that trained and untrained listeners, and even untrained
213 young children, are generally sensitive to stylistic aspects of music, and expertise enhances the
214 perception of style. However, most of the earlier work has only been able to speculate about the
215 types of cues used by trained and untrained listeners in such tasks. One reason is that the types of
216 tasks used in previous studies may not have been suitable to answer the primary question: How
217 do listeners categorize musical excerpts based on stylistic aspects? Studies involving the
218 measurement of implicit processes, such as those applied in the categorization of stimuli, should
219 use appropriate indirect or implicit investigation methods to obtain the best possible results. The
220 sorting task is designed for measuring implicit processes, such as those involved in most music-
221 related tasks. For instance, Brown (1981) found that trained and untrained participants agreed
222 less with their group performance when they had to pair melodies with descriptive words
223 provided by the researcher (explicit) versus when they did the matching task by providing their
224 own words (implicit). Similarly, Dibben (2001) found differences in participant responses
225 between nonverbal and verbal categorization tasks, wherein participants were more inclined to
226 group two sounds when they resembled each other acoustically in the implicit nonverbal task,
227 whereas they were more inclined to group them by their physical source in the explicit verbal
228 task. A second reason is that earlier tasks involving style perception and its analyses only
229 provided results obtained by averaging group responses, and there was no way to track

230 individual responses. This could be due to the fact that the statistical tools used to measure and
231 analyze the multidimensional aspects typically characteristic of human responses were not
232 sophisticated enough, and did not provide the possibility of projecting the results from various
233 perspectives (e.g., composer, participant expertise) into the same space, thus giving a
234 fragmentary account of those perspectives.

235 The only study we know of that has used a sorting task to categorize musical style (note
236 that here it was not composer's style but performer's style of playing) was conducted by Gingras,
237 Lagrandeur-Ponce, Giordano, and McAdams (2011). Experts and non-experts heard organ
238 excerpts represented as icons on a computer screen, which they had to sort into six groups based
239 on the performer's playing style. The excerpts were played by three award-winning and three
240 non-award-winning organists, who rendered two versions each of expressive and inexpressive
241 interpretation of the same piece. The results indicated that both trained and untrained participants
242 were able to accurately sort excerpts based on tempo, wherein faster excerpts were differentiated
243 from slower excerpts, and articulation (connectedness or separateness of notes in time), wherein
244 expressive performances were differentiated from inexpressive ones. Also, participants' sorting
245 was influenced by performer competence, in that they accurately differentiated between award-
246 winning versus other performers.

247 Although Gingras et al. (2011) successfully used the sorting task in a well-controlled
248 setting, in which participants heard versions of the same excerpt played in different ways by six
249 organists, an important issue that should be further studied is whether listeners can categorize
250 excerpts of different composers played by different performers based on stylistic features. Also,
251 Gingras et al. only used the constrained sort, which prompted us to investigate whether

252 participants would be able to categorize by composer's style when they are first told to sort freely
253 into as many groups as they see fit.

254 **Goals and Hypotheses**

255 The purpose of our study was to extend previous findings by examining the influence of
256 compositional style, type of sorting task (i.e., free vs. constrained), type of stimuli (MIDI vs.
257 natural), pianists' playing style, and listeners' music expertise on their ability to perceive stylistic
258 aspects in musical excerpts. We used both free and constrained sorting tasks and an updated
259 version of the statistical tool DiSTATIS, which had never before been used in studies pertaining
260 to music perception and cognition. One advantage of the sorting method is that the judgements
261 are more likely to reflect the multiple dimensions of the stimuli than would have been the case
262 when using the earlier paired comparison similarity judgements. Moreover, sorting tasks do not
263 involve any form of verbalization, thus tapping into the listeners' implicit knowledge. And such
264 a nonverbal approach facilitates the use and assessment of stylistic cues. Our study addressed the
265 following questions: (1) Are listeners able to sort brief melodies based on compositional style?
266 (2) If so, does the type of sorting task—free versus constrained—interact with music expertise in
267 influencing listeners' perception? Investigations involving other features of music, such as
268 emotion (e.g., Bigand, Vieillard, Madurell, Marozeau, & Dacquet, 2005; Bigand, Filipic, &
269 Lalitte, 2005), as well as research not involving music (e.g., Scott & Canter, 1997) have
270 measured listener responses in both sorting tasks while presenting the two tasks sequentially, free
271 sort followed by constrained sort. However, none of the sorting studies pertaining to musical
272 style that we have referenced have tested this, and so we decided to investigate what might
273 prompt differences, if any. Especially, we wanted to examine whether untrained participants in
274 particular could produce coherent categorizations for composer's style with the free sort. (3)

275 Does the type of stimuli—MIDI versus natural—influence the task? In contrast to the previous
276 studies, we compared listeners' perception of stylistic aspects in music between both MIDI and
277 natural stimuli excerpted from commercial recordings. (4) Also, will the performance of four
278 different pianists influence musical style perception? Unlike Dalla Bella and Peretz (2005), we
279 wanted to examine whether individual playing styles will influence participants' sorting choices
280 and the degree to which the different playing styles would affect listeners' perception of the
281 composers. (5) Finally, does the listener's music training influence the perception of stylistic
282 aspects of music?

283 Based on earlier studies (e.g., Crump, 2002; Dalla Bella & Peretz, 2005; Eastlund, 1992;
284 Eerola et al., 2001; Gardner, 1973; Gromko, 1993; Hargreaves & North, 1999; Miller, 1979;
285 Storino et al., 2007; Tekman & Hortaçsu, 2002; Thorisson, 1998; Tyler, 1946; Wedin, 1969), we
286 hypothesized that there would be an effect of musical period and compositional style, and that
287 listeners would identify greater stylistic differences among pieces from eras farther apart. That is,
288 participants would more distinctly categorize pieces by Bach and Beethoven versus those by
289 Bach and Mozart, or Mozart and Beethoven. Our second hypothesis was that, in general,
290 participants would be faster and more accurate in their perception of style in the constrained sort
291 when compared to the free sort. We based our prediction on the fact that participants completed
292 the constrained sort immediately following the free sort, which made them somewhat more
293 familiar with the excerpts. Also, the instructions were "clearer" with the constrained sort, where
294 we disclosed the actual number of composers. We expect that the results might indicate how far
295 increased familiarity and the change in instruction would change the result pattern. Our third
296 hypothesis was partly based on Gingras et al.'s (2011) findings, that participants would be more
297 accurate at perceiving the stylistic aspects in an expressive performance (natural stimuli) when

328 compared to a more mechanical one, since the natural stimuli have richer dynamics and tonal
329 qualities, thus facilitating perception. Our fourth hypothesis was also partly based on Gingras et
330 al.'s (2011) findings, that in Experiment 2, listeners would be influenced by the performance of
331 the pianists in their sorting of the pieces. Thus, composers' style would interact with pianists'
332 style wherein participants might incorrectly attribute pieces to other composers due to confusion
333 with the pianist's performance style. Our last hypothesis was that highly trained musicians would
334 be more accurate than untrained participants at perceiving the stylistic aspects. Previous research
335 has shown that professional musicians perceive musical structures differently from amateur
336 musicians (Dowling, 1986), with experts performing better than amateurs at a variety of musical
337 tasks (Krampe & Ericsson, 1996). However, several studies have also shown that untrained
338 participants are sensitive to underlying structural and affective patterns of music, and are able to
339 perform several musical tasks above chance levels (e.g., Bigand & Poulin-Charronnat, 2006;
340 Eastlund, 1992; Dalla Bella & Peretz, 2005; Gingras et al., 2011; Gromko, 1993; Wedin, 1969).
341 Thus, we also expect that untrained participants would display some style sensitivity.

342 **Framework for Compositional Style**

343 In this study, we used piano excerpts from three composers: Bach, Mozart, and
344 Beethoven. The composers' compositional styles are classified into three different epochs:
345 Bach's style is classified as baroque, Mozart's as classical, and Beethoven's as romantic. There
346 is general agreement among musicologists that Bach is a prototypical baroque composer who
347 played a very special role in the baroque period (Grout & Palisca, 1980). Nevertheless, his
348 compositions stood out in the baroque era due to their melodic, harmonic, and rhythmic
349 complexity. In general, though Bach occasionally shifted emotional tone in the middle of a
350 movement (especially in his cantatas), he typically followed the baroque style of maintaining a

321 constant emotional tone throughout a movement. Composers in the generation after Bach (e.g.,
322 Josef Haydn and Bach's son Carl Philipp Emmanuel), began to experiment with emotional shifts
323 within a movement, techniques that Mozart exploited during the classical period. In Beethoven,
324 the range and frequency of emotional shifts was expanded even further. Beethoven is right at the
325 start of the romantic approach, and his earlier works are usually viewed as transitional between
326 the classical and the romantic.

327 Especially in Experiment 2, we selected almost all of Beethoven's excerpts from his
328 romantic period. In accordance with the main features of musical style outlined above, we can
329 characterize these three styles in terms of the variability of the musical material along the
330 dimensions of pitch, time, and loudness, as well as of musical texture (dense vs. open and
331 transparent) and timbre. Timbre is not so much an issue in the present experiments because all
332 the excerpts are played on the piano, or in a piano timbre. And variations in loudness and tempo
333 are only an issue in Experiment 2 where we used actual performances of the pieces, in contrast to
334 the MIDI transcriptions of Experiment 1 which do not vary in loudness or in tempo. The
335 harmonic language differs among these styles—the way in which chord progressions and key
336 relationships are handled as the music develops in time (e.g., Zivic et al., 2013). Also,
337 Krumhansl (2015) cites results showing that there are differences in interval patterns between
338 baroque (largely scale wise) and romantic (largely arpeggios) styles. And there is a definite
339 change in the variability of emotional tone within an excerpt as we progress through the early
340 (baroque) and the middle (classical) 18th century and then on to the 19th (romantic).

341 The baroque style exemplified by Bach is characterized by rhythmic regularity as well as
342 relative stability of loudness, pitch, texture, and emotional tone within an excerpt. The harmonic
343 language is dense and complicated, taking surprising turns which are then resolved to achieve

344 expected ends. The texture typically consists of the interweaving of two or three separate
345 melodic lines in different pitch registers, which are clearly discernable. This is in contrast, for
346 example, to music that consists of a single melodic line accompanied by block chords in which
347 the individual pitches are not distinctly heard.

348 The classical style emerged from the baroque through the innovations of mid-
349 century composers, such as C.P.E. Bach, Haydn, and Mozart. The harmonic language becomes
350 simpler, often with a slower progression from chord to chord, but shifts of tonal center
351 (modulations) are often more abrupt, and signal a shift in emotional tone. Rhythmic organization
352 also becomes more irregular than in the baroque, accompanied by greater variation in loudness.
353 Textures are more varied, with dense as well as open textures, and often with a single melodic
354 line with chords or repetitive melodic figures outlining chords as accompaniment.

355 Beethoven, starting to write in the 1790s, shifted music into the romantic style. Here, the
356 tendencies apparent in classical music become accentuated. Especially for Beethoven (in contrast
357 to later romantics, such as Chopin, Schumann, and Brahms), the harmonic language becomes
358 even more simplified. Beethoven is sometimes inclined to emotional outbursts indicated by
359 abrupt changes in loudness, tempo, and texture. The range of pitches typically in use, expanded
360 somewhat in the classical compared with the baroque, is now widely expanded.

361 One aspect to consider with these three composers is that, spanning a century as they do,
362 their influence on each other is one-directional. Mozart was a dedicated admirer of Bach, and
363 from time to time there are unmistakable signs of Bach's influence. Mozart's String Trios, K.
364 404a, consists of his arrangements of Bach preludes and fugues, along with additional pieces to
365 go with them that he wrote in the same style. And the duet for the Two Armed Men in *The*
366 *Magic Flute* is definitely written in the style of Bach, which gives it a seriousness and solemnity

367 important to that scene in the opera. And Beethoven drew on both Bach and Mozart in his piano
368 music and string quartets, in which he includes passages where the interweaving of simultaneous
369 melodic lines is reminiscent of Bach. Beethoven was fond of Mozart's piano concertos,
370 especially Concerto No. 20 in d minor, which he often played in concerts and for which he wrote
371 a cadenza.

372 **Framework for Pianist Style**

373 Whereas in Experiment 1 we used MIDI transcriptions that simply reproduced the notes
374 on the printed page with no stylistically induced nuances in performance, such as variations in
375 loudness, tempo, and phrasing, in Experiment 2 we used excerpts from commercial recordings
376 played by four pianists: Claudio Arrau, Daniel Barenboim, Maria-João Pirès, and Sviatislav
377 Richter. We picked these pianists because they were among the relatively few pianists in the
378 middle and late 20th century who had recorded substantial amounts of the works of the three
379 composers. (Many pianists are known for concentrating on one composer, or several composers
380 in a similar stylistic period. Arthur Rubenstein, known for Beethoven and the later romantics, for
381 example, rarely if ever recorded Bach or Mozart, and Rosalyn Tureck, a Bach specialist, rarely
382 recorded Mozart or Beethoven.) We also selected them because their personal styles of playing
383 the piano differed systematically. Richter is widely regarded as presenting each composer, and
384 each piece, in its own terms, without imposing a particular personal imprint, but with
385 considerable emotional engagement (Villemin, 1999). His Bach is transparent and lucid, in that
386 the inner melodic lines are rarely obscured, but it is also forceful, as are his Mozart and
387 Beethoven. Arrau has been described as leaving his own imprint on the pieces he engages with
388 (Villemin, 1999), but his Bach is also transparent and his emotional engagement is very clear. In
389 contrast to Richter's playing which can often strike one as jagged and craggy, Arrau's is much

412 plain MIDI transcriptions available on the Internet. All excerpts were of pieces written for piano
413 or harpsichord, and we avoided pieces we judged to be relatively familiar, such as those found in
414 elementary piano books, like Bach's short minuets, Mozart's *Sonata in C, K. 545*, or
415 Beethoven's *Für Elise*. We produced .wav files of CD quality from MIDI files, as follows: The
416 MIDI files had been transcribed directly into MIDI form from the musical scores, with no
417 attention to nuances of dynamics, phrasing, or variations in tempo (e.g., files found at
418 <http://www.madore.org/~david/music/midi/>). The excerpts were converted into .wav files using
419 Cakewalk Professional version 4.0 using an acoustic piano voice. Each excerpt lasted for 9 to 10
420 s ($M = 9.67$ s). We linked the excerpts to audio icons arranged pseudorandomly on a PowerPoint
421 slide. Participants listened to the excerpts over good quality headphones.

422 **Sorting tasks: Free and constrained.** A sorting task is a simple but useful method to
423 examine implicit nonverbal processes, such as listeners' perception of stylistic aspects of the
424 melodies. Sorting tasks can reveal the underlying structure of a collection of items, in this case
425 musical excerpts. Similar to similarity-judgement and rating tasks, sorting tasks access implicitly
426 learned knowledge, in this case knowledge about stylistic aspects of melodies. However, sorting
427 tasks are considered to be more effective than judgement and rating tasks as they are less
428 strenuous on the participants, and can be used to compare experts and non-experts without
429 relying on either a specialized vocabulary or a quantitative response (Chollet, Valentin, & Abdi,
430 2014). For instance, we conducted interviews with participants in the pilot study about their
431 experiences with the sorting task, and they all uniformly reported that the task was fun and not at
432 all tiring. Some researchers have indicated that untrained participants and especially children
433 have difficulty in verbalizing their perceptual responses to art forms (e.g., Gardner, 1973). Other
434 researchers have shown that participants' performance change in a verbal versus nonverbal

435 musical task (e.g., Brown, 1981; Dibben, 2001). An added theoretical advantage for a nonverbal
436 sorting task could be that participants are able to use their own concepts for categorizing stimuli
437 especially in a free sort (e.g., Scott & Canter, 1997), and are completely in control of the
438 experiment in terms of its pace and time limit, so they probably find the task less taxing. As in
439 Bigand, Vieillard, Madurell, Marozeau, and Dacquet (2005), Bigand, Filipic, and Lalitte (2005),
440 and Scott and Canter (1997), the order of the two sorting tasks, free and constrained, could not be
441 alternated as the purpose of the free sort was to have participants categorize the excerpts without
442 specific direction, whereas the purpose of the constrained sort was to re-categorize the same
443 excerpts based on specific instructions. Sorting tasks are commonly employed in studies on
444 sensory perception, such as food preference and quality. As far as we know, only one study (i.e.,
445 Gingras et al., 2011) has employed this method to assess listeners' perception of musical style,
446 though listeners regularly apply sorting methods in their everyday musical and nonmusical
447 activities.

448 **Procedure.** In our experiment, we adapted the methodology used in Bigand, Vieillard, et
449 al. (2005), and Bigand, Filipic, and Lalitte, (2005). We instructed participants to listen to each
450 excerpt by clicking on its icon, and then to sort the icons into clusters based on their perceived
451 similarity to each other—in particular based on whether they might have been written by the
452 same composer. While sorting, participants could play each excerpt in any order they wished and
453 as many times as they wanted similar to the methodology used by Bigand et al. and Gingras et al.
454 (2011), and especially as the stimuli were presented to them in a random order with each task.
455 However, we did not register the number of times participants heard each stimulus nor the order
456 in which the stimuli were heard. Participants completed two types of sorting tasks sequentially:
457 free sorting and constrained sorting. In the free sorting task, participants sorted the 21 excerpts

458 into as many clusters as they thought necessary, with the constraints that there should be at least
459 two clusters, and that each cluster should contain at least two excerpts. In the constrained sorting
460 task, participants were required to sort the excerpts into three clusters only as we had excerpts
461 from three composers; this gave them more direction and should have helped them in sorting.
462 The whole task took approximately 20 to 40 min to complete, depending on how often the
463 participant listened to the various excerpts.

464 **Data Analysis.** We recoded each participant's sorting data as a distance matrix. Excerpts
465 sorted together were assigned a distance of 0, whereas excerpts sorted into different groups were
466 assigned a distance of 1. To analyze the perceived differences among the excerpts, we then used
467 an updated version of DiSTATIS (Abdi, Williams, Valentin, & Bennani-Dosse, 2012).

468 DiSTATIS is a generalization of two multivariate methods: metric multidimensional scaling
469 (MDS; Abdi, 2007b), a method for analyzing a single distance matrix, and STATIS, a method
470 for executing multi-table principal component analysis (PCA; Abdi et al., 2012; Abdi &
471 Williams, 2010). DiSTATIS is commonly used to assess multiple distance matrices, such as data
472 from sorting tasks (Abdi, Valentin, Chollet, & Chrea, 2007; Abdi, 2007a), wherein each
473 participant produces a distance matrix. Here our application of DiSTATIS relies on *a priori*
474 knowledge, namely the fact that we used excerpts from exactly three composers (and in
475 Experiment 2, four pianists).

476 In DiSTATIS, participants' distance matrices are double-centered, normalized, integrated
477 (i.e., combined), and decomposed to give a factor map. To double-center the matrices (Abdi,
478 2007b), a distance matrix is converted to a covariance matrix centered on the origin. In this way,
479 double-centering brings disparate matrices to the same center (similar to centering as in
480 calculating *z* scores). Double-centered matrices are normalized in the style of multiple factor

481 analysis (Abdi, Williams, & Valentin, 2013), where each double-centered matrix is divided by its
482 first eigenvalue so that the scales of the tables are comparable. These double-centered and
483 normalized tables are then subjected to an analysis of between-table similarity, called R_V analysis
484 (Abdi, 2010), in order to identify typical and atypical tables. The R_V analysis provides a set of
485 table weights, such that atypical tables receive small weights. The weighted average of these
486 tables gives the best possible single representation of all the tables, called the compromise table
487 (Abdi et al., 2012). Finally, the compromise table is decomposed by PCA to generate
488 components. Thus, DiSTATIS reveals the best possible single representation of the perceived
489 relationships among the stimuli.

490 The advantage of using DiSTATIS is that, unlike MDS and PCA, it retains the
491 information provided by the pattern of each participant's responses, but like MDS and PCA,
492 DiSTATIS produces new variables, called components (also called dimensions, factors, or
493 principal axes). Components are ordered by strength and are mutually orthogonal. That is, the
494 first component explains the maximum possible variance, and the subsequent components
495 explain the maximum remaining variance under the constraint that each subsequent component is
496 orthogonal to all prior components. The coordinates of the stimuli on the components are called
497 factor scores.

498 For ease of visualization, typically two components are plotted in what is called a
499 component map. On this map, observations are interpreted by their distances from each other and
500 their positions on the components. Observations near each other are similar. An observation that
501 has a large factor score on a given component contributes much variance to that component.
502 Each component may reflect an effect measured along that dimension, which may relate to a
503 perceived difference between the observations (e.g., staccato vs. legato). Thus, two observations

504 on the same side of a component are perceived as similar on that dimension, whereas
505 observations on opposite sides of a component are perceived as different on that dimension. In
506 the figures, dots representing excerpts are color-coded by composer, and a square box in the
507 appropriate color indicates each composer's average position.

508 We also performed inference tests in the form of nonparametric bootstrap resampling to
509 test the stability of differences between groups. We tested the differences among the three groups
510 of composers, and also among three levels of music training of the participants: 11 untrained
511 participants, 10 moderately trained musicians (1 to 4 years of training, $M = 2.60$ years), and 18
512 highly trained musicians (5 or more years of training, $M = 10.28$ years). Previous studies have
513 shown that people with 5 or more years of formal music training perform differently on musical
514 tasks than those with less than 5 years of training or those with no training at all. For example,
515 Dowling (1986), and Dowling and Bartlett (1981) showed strong differences in performance
516 between people with average of 5 years of music lessons than those without any. Bootstrap
517 resampling consists of resampling participants within groups with replacement (DiCiccio &
518 Efron, 1996), a procedure intended to simulate sampling from the population of individuals from
519 which the participants are drawn. Bootstrap samples are collected repeatedly (here, 1,000 times)
520 to generate successive distributions of the groups. For each group, the most extreme 5%
521 bootstrap-sampled means are removed, leaving a peeled convex hull that contains 95% of the
522 bootstrap-sampled means, giving a 95% bootstrap confidence interval. For visualization, a
523 smoothed ellipse is fitted around the convex hull, and so is slightly more conservative than the
524 convex hull itself. We conducted the analyses in R (version 2.15.2; R Core Team, 2012),
525 adapting the DistatisR (Beaton, Chin-Fatt, & Abdi, 2014a; Beaton, Chin-Fatt, & Abdi, 2014b)
526 and the MExPosition (Chin-Fatt, Beaton, & Abdi, 2013) packages to that use.

527 Results

528 We conducted DiSTATIS analyses on the data from Experiments 1 and 2 separately for
529 the free and constrained sorting tasks. Table 1 shows the percent of variance explained by the
530 first four components in each of the four overall analyses in which the sorting was based on
531 composers. These components explain between 5.28 and 21.47% of the variance in the four
532 analyses.

533 Free Sorting.

534 *Composers.* Figure 1 shows that Components 1, 2, and 3 captured the effects of
535 composer. Component 1 differentiated Beethoven from the other two composers. To a lesser
536 extent Component 2 differentiated Mozart from the other two. Component 3 differentiated Bach
537 from Mozart and Beethoven, whereas for Component 4 there were no apparent differences due to
538 composer.

539 *Music Training.* Figure 2 shows the results of the R_V analysis for sorting patterns
540 produced by the participants. Here, each dot corresponds to a participant. Participants were
541 color-coded according to level of music training: highly trained, moderately trained, and
542 untrained. Component 1 displayed a non-significant effect in which highly trained musicians
543 were separated from the others. Component 2 indicated a separation between moderately trained
544 musicians and untrained participants, but with highly trained musicians in between. Subsequent
545 components did not reveal between-group effects.

546 Constrained Sorting.

547 *Composers.* In Figure 3, Components 1, 2, and 4 showed that pieces by Beethoven were
548 clearly distinguished from those of the other composers. Component 3 differentiated Bach from
549 the other composers.

550 **Music Training.** Figure 4 shows the results of the R_V analysis. No effects of music
551 training were found; and the three groups behaved more similarly with constrained sorting than
552 they did with free sorting.

553 **Experiment 2: Natural stimuli**

554 **Method**

555 **Participants.** Thirty-seven participants, mean age 23.27 years (range = 17 to 50 years),
556 took part in Experiment 2. Ten participants reported that they had no music training whereas the
557 remaining 27 participants reported that they had between 1 and 15 years ($M = 4.89$ years) of
558 formal music training.

559 **Stimuli.** Stimuli consisted of 36 newly selected excerpts from commercial CD
560 recordings: 12 pieces each by Bach, Mozart, and Beethoven. Each of four pianists—Arrau,
561 Barenboim, Pirès, and Richter—played three different pieces by each composer (see Appendix
562 B). This enabled us to assess the constancy of a composer’s place in the sorting patterns across
563 varied pianists, and the degree to which differences among the pianists affected sorting. As in
564 Experiment 1, we avoided relatively familiar works. We were constrained by the selection of
565 works that the particular pianists had recorded. For example, Richter had mainly recorded Bach’s
566 *Wohltemperierte Klavier*, whereas Arrau had mainly recorded partitas and suites. In contrast to
567 Experiment 1, these excerpts exhibited all the nuances of phrasing and dynamics characteristic of
568 musical performances. Each excerpt was at least 9 s in length, and ended at a musically coherent
569 place, so that they varied in length from 9 to 15 s ($M = 11.64$ s). Presentation of stimuli was the
570 same as in Experiment 1.

571 **Procedure and Data Analysis.** The procedure and data analysis were the same as those
572 of Experiment 1 except for the following differences: The total duration of the task was

573 approximately 30 to 45 min, depending on the participant. And the groupings addressed by the
574 DiSTATIS nonparametric bootstrap resampling analyses included contrasts among the four
575 pianists as well as among composers. Categorization of participants' expertise was the same as in
576 Experiment 1, with 10 untrained participants, 17 moderately trained musicians ($M = 2.06$ years),
577 and 10 highly trained musicians ($M = 9.70$ years).

578 **Results**

579 **Free Sorting.**

580 **Composers.** In Figure 5, Components 1 and 2 differentiated Mozart from Beethoven,
581 with Bach's excerpts clustered near the origin. Component 3 differentiated Bach from
582 Beethoven, and Component 4 differentiated Mozart from the other two.

583 **Pianists.** Figure 6 shows the results regarding pianists. Component 1 differentiated
584 Richter from Pirès. Component 2 differentiated Richter and Pirès from Barenboim, with Arrau in
585 the middle. Component 3 differentiated between Richter and Barenboim with Arrau and Pirès in
586 the middle. Arrau was consistently positioned near the origin. Note that Figures 5a and 6a
587 suggest a connection between composer and pianist, such that Barenboim appears to be
588 definitely associated with Beethoven, with Pirès and Richter more associated with Mozart, and
589 Arrau appears at the origin along with Bach.

590 **Music Training.** Figure 7 shows the results of the R_V analysis. Component 1 indicated an
591 effect of musical experience, significantly separating low and high levels of musical training
592 with moderate levels in between. There were no other clear effects.

593 **Constrained Sorting.**

594 **Composers.** In Figure 8, Components 1, 2, and 4 distinguished between Beethoven and
595 Mozart, whereas Component 3 differentiated Bach from the other two.

596 *Pianists.* Figure 9 shows the results in terms of pianists. Richter and Barenboim were
597 consistently perceived as distinct. Component 2 differentiated Pirès from Barenboim.
598 Components 3 and 4 taken together distinguished Barenboim and Arrau from Richter and Pirès.
599 Note that Figures 8a and 9a show a relationship between composer and pianist, similar to that
600 seen in Figures 5a and 6a.

601 *Music Training.* Figure 10 shows the results of the R_v analysis. Component 1 displayed a
602 non-significant difference indicating a weak effect of musical training. There were no other
603 effects.

604 **Discussion**

605 In considering these results, let us first look at the contrast between free sorting and
606 constrained sorting. In general, constrained sorting produced greater agreement among the
607 listeners than free sorting (which was done first), as shown by the amount of variance explained
608 by the successive factors in the DiSTATIS solutions (see Table 1). Especially in Experiment 2,
609 the gain attributable to constrained sorting is substantial. In both experiments, the total amount of
610 variance explained for constrained sorting by the first four factors is around 50 %, compared
611 with about 40 % in Experiment 1 and about 30 % in Experiment 2 for free sorting. Constraining
612 the sorting to just three categories forced listeners to make difficult choices of whether to put
613 excerpts in the same cluster, which they had perhaps avoided in the free sort by creating more
614 categories. And those choices led to greater consistency and agreement among the listeners in
615 their categorization of style. This increase in consistency was accompanied by greater
616 convergence among the groups of listeners with different amounts of musical training, as seen in
617 going from Figure 2 to Figure 4 for Experiment 1, and from Figure 7 to Figure 10 for
618 Experiment 2. These results suggest that in constrained sorting, the untrained and the moderately

619 trained groups appear to be using much the same features for making decisions about
620 compositional style. And there is considerable overlap between the features they use and the
621 features used by the more highly trained groups. In general, in this regard these results agree with
622 those of Dalla Bella and Peretz (2005), Eastlund (1992), Gingras et al. (2011), Gromko (1993),
623 and Wedin (1969). These results also concur with those of Brown (1981) and Dibben (2001), in
624 that untrained and trained listeners perform similarly in nonverbal music tasks. In regard to
625 Miller's (1979) finding that untrained listeners tend to rely more on affective qualities of the
626 excerpts, we note that the convergence across training levels was more emphatic in Experiment
627 2, where those affective qualities were more evident in the naturalistic excerpts, than in
628 Experiment 1, where they were largely absent.

629 Since the constrained sorts were more coherent than the free sorts, in what follows we
630 will concentrate on them. In Experiment 1, Figure 3a shows that the first component tends to
631 separate the three composers according to their historical order: Bach, Mozart, and Beethoven.
632 This is in agreement with the results of Dalla Bella and Peretz (2005) who suggested that this
633 categorization was largely driven by the increase in rhythmic freedom as style developed from
634 the baroque through the classical to the romantic. Such an increase in rhythmic freedom involves
635 features that would be quite evident in the MIDI versions of Experiment 1, so this interpretation
636 strikes us as entirely appropriate. The second component in Figure 3a appears to contrast Bach
637 and Mozart with Beethoven. Among the readily available features in the MIDI excerpts,
638 harmonic complexity suggests itself as underlying this contrast: Bach and Mozart are notably
639 more complex in their harmonic progressions than Beethoven, especially the relatively early
640 Beethoven represented in Experiment 1 (see Appendix A). (This local trend runs counter to the
641 more general historical trend noted by Dalla Bella and Peretz of an increasingly freer use of the

642 tonal system over the last three centuries.) The third component (Figure 3b) contrasts Mozart and
643 Beethoven with Bach, and may have to do with constancy of texture. As noted above, Bach's
644 writing typically involved the simultaneous presentation of two or three separate melodic lines in
645 a texture that remains generally constant throughout an excerpt, and this textural consistency is
646 obvious in these MIDI excerpts. Mozart and Beethoven, in contrast, shift their textures often,
647 between few versus many simultaneous notes, and between pitch regions, and those shifts are
648 also obvious in the MIDI transcriptions. Component 4 (Figure 3b) again contrasts Bach and
649 Mozart with Beethoven, but we do not venture an interpretation.

650 Turning to the naturalistic excerpts of Experiment 2, we see in Component 1 (Figure 8a)
651 an ordering of Beethoven-Bach-Mozart. With the live pianists we think this reflects differences
652 in the forcefulness of the performances, involving dynamic (loudness) contrasts. Beethoven uses
653 the greatest dynamic contrasts, and with these pianists Bach comes a close second, whereas
654 Mozart is more reserved and delicate. Component 2 (Figure 8a) appears to reflect large-scale
655 rhythmic unpredictability, in which the less predictable Beethoven is contrasted with the more
656 predictable Bach and Mozart. This contrast was accentuated in the live performances because the
657 pianists tended to give dynamic emphasis to Beethoven's rhythmic surprises, which led to a
658 different result here than in Experiment 1 (see Component 1 in Figure 3a) where no such
659 emphasis could occur. As a result, the three composers do not line up in historical order on what
660 we are thinking of as a dimension of rhythmic complexity as they did in Experiment 1 and in
661 Dalla Bella and Peretz (2005). Component 3 (Figure 8b) comes close to putting them in
662 historical order, though Mozart and Beethoven overlap to a considerable degree. We think this
663 dimension can be attributed to variability of texture, similar to Component 3 of Experiment 1
664 (Figure 3b). Bach's pieces tend to stick with the same relatively closed texture for long periods

665 of time, in contrast to those of Mozart and Beethoven, who often shift the texture in density and
666 pitch range. We interpret Component 4 (Figure 8b) as concerned with emotional engagement.
667 Mozart (probably now more than in his own time) tends to be heard as elegant and above the
668 fray, whereas Bach, and to an even greater extent Beethoven, tend to be heard as passionate and
669 emotionally engaged. For Bach, this is especially true in performances by the pianists
670 represented here (especially Barenboim, Arrau, and Richter), in contrast to a number of pianists
671 who specialize in Bach, such as Glenn Gould and Rosalyn Tureck.

672 We now turn to the constrained sorts of Experiment 2 viewed in terms of the pianists
673 (Figures 9a & 9b). Keep in mind that the solution that underlies these figures is the same as the
674 solution in Figure 8; that is, all the individual points pertaining to excerpts are the same, but now
675 the means (“barycenters”) are calculated by grouping each pianist’s points together, rather than
676 each composer’s. So, Component 1 (Figure 9a) appears to indicate affinities between the pianists
677 and particular composers: Barenboim with Beethoven, Arrau and Pirès with Bach, and Richter
678 with Mozart. This last pairing is somewhat of a surprise, as in his career Richter was more
679 typically associated with Beethoven and Bach than with Mozart. On the other hand, as Villemain
680 (1999) noted, Richter was known for adapting his style to that of the composer he was playing,
681 and so among the pianists here he may have been the best fit for Mozart. His Mozart in these
682 excerpts was certainly among the most expressive performances of them. Component 2 (Figure
683 9a) may concern overall heaviness of the texture, ranging from the relatively dark and heavy
684 piano sound of the early Barenboim, to a moderately heavy sound of Arrau and Richter, to the
685 very light sound of Pirès. This order parallels the progression from Bach through Beethoven to
686 Mozart in Component 2 for composers (Figure 8a). Component 3 (Figure 9b) represents clarity
687 of texture: Barenboim and Arrau (denser) versus Richter and Pirès (clear and lucid). (Component

688 3 for composers (Figure 8b) contrasted the relatively dense Bach with the more open Mozart and
689 Beethoven.) And we do not venture to interpret Component 4 (Figure 9b).

690 One of the primary goals of this study was to verify the effectiveness of a sorting task and
691 its analysis using DiSTATIS in musical style perception. The results showed that the sorting task
692 could be successfully used to ascertain listeners' implicit knowledge of stylistic aspects,
693 especially for untrained listeners. Many participants reported that they "had fun" sorting the
694 excerpts, and that this task seemed less strenuous on them. Both music experts and non-experts
695 performed similarly especially since the task did not rely on using technical vocabulary or any
696 form of verbalization or quantification. On the other hand, knowing the basis of categorization of
697 the melodies might help researchers understand the exact nature of cues (i.e., high-level or low-
698 level) that each participant uses, and also to ascertain whether music training would influence the
699 type of cues that listeners perceive. In a future study, researchers could ask participants to label
700 each group of melodies after they complete both the free and constrained sorting tasks. Another
701 reason for the untrained participants' competent performance on this task could be the use of
702 excerpts from actual artistic performances, which contain a repository of cues pertaining to
703 dynamics, texture, and so forth, not present in the MIDI versions of Experiment 1. Our study
704 clearly showed that years of mere passive listening could facilitate the perception of such cues.
705 One limitation of this study was that we did not assess the familiarity of our participants with
706 each stimulus, and thus, we cannot estimate whether veridical knowledge (i.e., piece-specific
707 information) helped the trained listeners in doing the task. Nevertheless, we did use relatively
708 unfamiliar excerpts (see Appendices A and B), and, most importantly, we did not see much
709 differences in the relative performance of the untrained and the two trained groups. Also, all
710 participants performed the free sort first followed by the constrained sort, as by definition it is

711 impossible to counterbalance the order of presentation (see also Bigand, Vieillard, et al., 2005;
712 Bigand, Filipic, & Lalitte, 2005; Scott & Canter, 1997, using the same order). This meant that all
713 participants doing the constrained sort were more familiar with the excerpts than in the free sort,
714 thus potentially contributing to more coherent and converging responses in the second task.
715 However, there is no confound here since all participants did the two tasks in the same sequence.
716 Moreover, we re-randomized the order of the excerpts in the second task, so that participants had
717 to re-categorize the excerpts based on the “new” constraints provided by the experimenter. A
718 second limitation of the study is that we did not track the number of times participants heard
719 each excerpt. For instance, Gingras et al. (2011), using the same sorting paradigm, found that the
720 total number of times participants listened to each excerpt correlated significantly with their
721 categorization accuracy. Also, an influence of musical expertise on problem-solving behavior in
722 a musical puzzle task was reported by Tillmann, Bigand, and Madurell (1998); in particular,
723 trained participants listened more often to the musical puzzle parts, but less often to the entire
724 musical piece than did untrained participants. These overall findings convergently show that
725 music experts tended to listen to the stimuli (or the parts of the puzzle individually) more often
726 than the novices, which probably enhanced the experts’ performance in the task. Building on
727 these findings and our study, a future sorting experiment could investigate such a relationship
728 between musical expertise and problem solving or perceptual strategies further. Future studies
729 could also address if this task would be successful in discerning subtler and more nuanced
730 aspects of musical style. For instance, would trained and untrained participants be able to sort
731 melodies based on early versus late Beethoven’s compositional style? Finally, an important
732 follow-up experiment would be to investigate the effectiveness of this task when applied in a
733 cross-cultural musical style perception study with expertise and familiarity as factors.

734 **References**

- 735 Abdi, H. (2007a). Distance. In N. J. Salkind (Ed.), *Encyclopedia of measurement and statistics*
736 (pp. 280-284). Thousand Oaks, CA: Sage. doi: 10.4135/9781412952644.n142
- 737 Abdi, H. (2007b). Metric multidimensional scaling (MDS): Analyzing distance matrices. In N. J.
738 Salkind (Ed.), *Encyclopedia of measurement and statistics* (pp. 598–605). Thousand
739 Oaks, CA: Sage. doi: 10.4135/9781412952644.n279
- 740 Abdi, H. (2010). Congruence: Congruence coefficient, R_V coefficient, and Mantel coefficient. In
741 N. J. Salkind (Ed.), *Encyclopedia of research design* (pp. 222-229). Thousand Oaks, CA:
742 Sage. doi: 10.4135/9781412961288.n71
- 743 Abdi, H., Valentin, D., Chollet, S., & Chrea, C. (2007). Analyzing assessors and products in
744 sorting tasks: DiSTATIS, theory and applications. *Food Quality & Preference*, 18(4),
745 627–640. doi:10.1016/j.foodqual.2006.09.003
- 746 Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary*
747 *Reviews: Computational Statistics*, 2(4), 433–459. doi:10.1002/wics.101
- 748 Abdi, H., Williams, L. J., & Valentin, D. (2013). Multiple factor analysis: Principal component
749 analysis for multi-table and multi-block data sets. *Wiley Interdisciplinary Reviews:*
750 *Computational Statistics*, 5(2), 149–179. doi:10.1002/wics.1246
- 751 Abdi, H., Williams, L. J., Valentin, D., & Bennani-Dosse, M. (2012). STATIS and DiSTATIS:
752 Optimum multitable principal component analysis and three way metric
753 multidimensional scaling. *Wiley Interdisciplinary Reviews: Computational Statistics*,
754 4(2), 124–167. doi:10.1002/wics.198

- 755 Agres, K., Abdallah, S., & Pearce, M. (2018). Information-theoretic properties of auditory
756 sequences dynamically influence expectation and memory. *Cognitive Science*, *42*, 43-76.
757 doi: 10.1111/cogs.12477
- 758 Atalay, B. V., & Placek, R. (1997). Machine versus human: Responding to the task of identifying
759 eras for selected keyboard pieces. In A. Gabrielsson (Ed.), *Proceedings of the Third*
760 *Triennial ESCOM Conference* (pp. 521-526). Uppsala, Sweden: Uppsala University.
- 761 Beaton, D., Chin-Fatt, C. R., & Abdi, H. (2014a). An ExPosition of multivariate analysis with
762 the singular value decomposition in R. *Computational Statistics & Data Analysis*, *72*,
763 176–189. doi:10.1016/j.csda.2013.11.006
- 764 Beaton, D., Chin-Fatt, C. R., & Abdi, H. (2014b). Package “DistatisR.”
- 765 Bigand, E., Filipic, S., & Lalitte, P. (2005). The time course of emotional responses to music.
766 *Annals of the New York Academy of Sciences*, *1060*(1), 429-437. doi:
767 10.1196/annals.1360.036
- 768 Bigand, E., & Poulin-Charronnat, B. (2006). Are we “experienced listeners”? A review of the
769 musical capacities that do not depend on formal musical training. *Cognition*, *100*(1), 100-
770 130. doi:10.1016/j.cognition.2005.11.007
- 771 Bigand, E., Vieillard, S., Madurell, F., Marozeau, J., & Dacquet, A. (2005). Multidimensional
772 scaling of emotional responses to music: The effect of musical expertise and of the
773 duration of the excerpts. *Cognition & Emotion*, *19*(8), 1113-1139. doi:
774 10.1080/02699930500204250
- 775 Brown, R. (1981). Music and language. In *Documentary report of the Ann Arbor Symposium:*
776 *National symposium on the applications of psychology to the teaching and learning of*
777 *music*. Reston, VA: Music Educators National Conference.

- 778 Chin-Fatt, C. R., Beaton, D., & Abdi, H. (2013). Package “MExPosition.”
- 779 Chollet, S., Valentin, D., & Abdi, H. (2014). Free sorting task. In P.V. Tomasco & G. Ares
780 (Eds.), *Novel techniques in sensory characterization and consumer profiling* (pp. 207-
781 227). Boca Raton, FL: Taylor & Francis.
- 782 Cope, D. (1991). *Computers and musical style*. Madison, WI: A-R Editions.
- 783 Crump, M. (2002). *A principal components approach to the perception of musical style*
784 (Unpublished honors thesis). University of Lethbridge, Canada.
- 785 Dalla Bella S., & Peretz, I. (2005). Differentiation of classical music requires little learning but
786 rhythm. *Cognition*, 96, B65-B78. doi: 10.1016/j.cognition.2004.12.005
- 787 DiCiccio, T. J., & Efron, B. (1996). Bootstrap confidence intervals. *Statistical Science*, 11(3),
788 189–212. Retrieved from <https://www.jstor.org/stable/2246110>
- 789 Dibben, N. (2001). What do we hear, when we hear music?: Music perception and music
790 material. *Musicae Scientiae*, 5(2), 161-194. doi: 10.1177/102986490100500203
- 791 Dowling, W. J. (1986). Context effects on melody recognition: Scale-step versus interval
792 representations. *Music Perception*, 3(3), 281-296. doi:10.2307/40285338
- 793 Dowling, W. J., & Bartlett, J. C. (1981). The importance of interval information in long-term
794 memory for melodies. *Psychomusicology*, 1(1), 30-49. doi: 10.1037/h0094275
- 795 Eastlund, J. O. (1992). A multidimensional scaling analysis of musical style. *Journal of Research*
796 *in Music Education*, 40(3), 204-215. doi: 10.2307/3345682
- 797 Eerola, T., Järvinen, T., Louhivuori, J., & Toiviainen, P. (2001). Statistical features and
798 perceived similarity of folk melodies. *Music Perception*, 18(3), 275-296. doi:
799 10.1525/mp.2001.18.3.275

- 800 Gardner, H. (1973). Children's sensitivity to musical styles. *Merrill-Palmer Quarterly of*
801 *Behavior and Development*, 19(1), 67-77. Retrieved from
802 <https://www.jstor.org/stable/23083794>
- 803 Gingras, B., Lagrandeur-Ponce, T., Giordano, B. L., & McAdams, S. (2011). Perceiving musical
804 individuality: Performer identification is dependent on performer expertise and
805 expressiveness, but not on listener expertise. *Perception*, 40, 1206-1220. doi:
806 10.1068/p6891
- 807 Gromko, J. E. (1993). Perceptual differences between expert and novice music listeners: A
808 multidimensional scaling analysis. *Psychology of Music*, 21(1), 34-47. doi:
809 10.1177/030573569302100103
- 810 Grout, D. J., & Palisca, C. V. (1980). *A history of western music* (3rd ed.). New York, NY: W. W.
811 Norton & Company, Inc.
- 812 Gundlach, R. H. (1932). A quantitative analysis of Indian music. *The American Journal of*
813 *Psychology*, 44(1), 133-145. doi: 10.2307/1414960
- 814 Hargreaves, D. J., & North, A. C. (1999). Developing concepts of musical style. *Musicae*
815 *Scientiae*, 3(2), 193-216. doi: 10.1177/102986499900300203
- 816 Hevner, K. (1936). Experimental studies of the elements of expression in music. *The American*
817 *Journal of Psychology*, 48(2), 246-268. doi: 10.2307/1415746
- 818 Järvinen, T., Toiviainen, P., & Louhivuori, J. (1999). Classification and categorization of
819 musical styles with statistical analysis and self-organizing maps. In A. Patrizio, G. A.
820 Wiggins, & H. Pain (Eds.), *Proceedings of the AISB '99 Symposium on Musical Creativity*
821 (pp. 54-57). Edinburgh, Scotland: The Society for the Study of Artificial Intelligence and
822 Simulation of Behavior.

- 823 Krumhansl, C. L. (2015). Statistics, structure, and style in music. *Music Perception*, 33(1), 20-
824 31. doi: 10.1525/mp.2015.33.1.20
- 825 Krampe, R. Th., & Ericsson, K. A. (1996). Maintaining excellence: Deliberate practice and elite
826 performance in young and older pianists. *Journal of Experimental Psychology: General*,
827 125, 331-359. doi:10.1037/0096-3445.125.4.331
- 828 Meyer, L. B. (1973). *Explaining music: Essays and explorations*. Berkeley, CA: University of
829 California Press.
- 830 Meyer, L. B. (1989). *Style and music: Theory, history, and ideology*. Philadelphia, PA:
831 University of Pennsylvania Press.
- 832 Miller, R. F. (1979). *An analysis of musical perception through multidimensional scaling*
833 (Doctoral dissertation). Retrieved from ProQuest dissertations & theses global. (Order
834 No. 8004230)
- 835 Morrison, S. J., & Yeh, C. S. (1999). Preference responses and use of written descriptors among
836 music and nonmusic majors in the United States, Hong Kong, and the People's Republic
837 of China. *Journal of Research in Music Education*, 47(1), 5-17. doi: 10.2307/3345824
- 838 Pascall, R. J. (1980). Style. In S. Stanley (Ed.), *The new grove dictionary of music and musicians*
839 (pp. 316-321). London, UK: Macmillan.
- 840 R Core Team (2012). *R: A language and environment for statistical computing*. Vienna, Austria:
841 R Foundation for Statistical Computing.
- 842 Raman, R., & Dowling, W. J. (2016). Real-time probing of modulations in South Indian classical
843 (Carnātic) music by Indian and Western musicians. *Music Perception*, 33(3), 367-393.
844 doi: 10.1525/mp.2016.33.03.367

- 845 Raman, R., & Dowling, W. J. (2017). Perception of modulations in South Indian classical
846 (Carnātic) music by student and teacher musicians: A cross-cultural study. *Music*
847 *Perception*, 34(4), 424-437. doi: 10.1525/mp.2017.34.4.424
- 848 Scott, M. J., & Canter, D. V. (1997). Picture or place? A multiple sorting study of landscape.
849 *Journal of Environmental Psychology*, 17(4), 263-281. doi: 10.1006/jevp.1997.0068
- 850 Smaill, A., & Westhead, M. D. (1993). Automatic characterization of musical style. In M. Smith,
851 A. Smaill, & G. A. Wiggins (Eds.), *Music education: An artificial intelligence approach*
852 (pp. 157-170). Berlin, Germany: Springer-Verlag.
- 853 Storino, M., Dalmonte, R., & Baroni, M. (2007). An investigation on the perception of musical
854 style. *Music Perception*, 24(5), 417-432. doi: 10.1525/mp.2007.24.5.417
- 855 Tekman, H. G., & Hortaçsu, N. (2002). Aspects of stylistic knowledge: What are different styles
856 like and why do we listen to them? *Psychology of Music*, 30, 28-47. doi:
857 10.1177/0305735602301005
- 858 Tillmann, B., Abdi, H., & Dowling, W. J. (2004). Musical style perception by a linear auto-
859 associator model and human listeners. In S. D. Lipscomb, R. Ashley, R. O. Gjerdingen,
860 & P. Webster (Eds.), *Proceedings of the 8th International Conference on Music*
861 *Perception & Cognition (ICMPC)* (pp. 189-191). Adelaide, Australia: Causal
862 Productions.
- 863 Tillmann, B., Bigand, E., & Madurell, F. (1998). Local versus global processing of harmonic
864 cadences in the solution of musical puzzles, *Psychological Research*, 61(3), 157-174. doi:
865 10.1007/s004260050022

- 866 Thorisson, T. (1998). Comparison of novice listeners' similarity judgements and style
867 categorisation of classic and romantic piano exemplars. *Psychology of Music*, 26(2), 186-
868 196. doi: 10.1177/0305735698262006
- 869 Tyler, L. E. (1946). An exploratory study of discrimination of composer style. *The Journal of*
870 *General Psychology*, 34(2), 153-163. doi: 10.1080/00221309.1946.10544531
- 871 Vignal, M. (1987). *Le dictionnaire de la musique*. Paris, France: Larousse.
- 872 Villemin, S. (1999). *Les grands pianistes*. Geneva, Switzerland: Georg.
- 873 Watt, R. J., & Ash, R. L. (1998). A psychological investigation of meaning in music. *Musicae*
874 *Scientiae*, 2(1), 33-53. doi: 10.1177/102986499800200103
- 875 Wedin, L. (1969). Dimension analysis of the perception of musical style. *Scandinavian Journal*
876 *of Psychology*, 10(1), 97-108. doi: 10.1111/j.1467-9450.1969.tb00015.x
- 877 Zivic, P. H. R., Shifres, F., & Cecci, G. A. (2013). Perceptual basis of evolving Western musical
878 styles. *Proceedings of the National Academy of Sciences*, 110(24), 10034-10038.
879 Retrieved from <http://www.jstor.org/stable/42706125>
- 880
- 881
- 882
- 883
- 884
- 885
- 886
- 887
- 888

889

Appendix A

890

Experiment 1: MIDI Stimuli

No.	Composer	Key	Title
1	Bach	A	English Suite No. 1, Guigue 806
2		B ^b	Partitas No. 1, BWV 825
3		C	Three-Part Invention, BWV 787
4		C-minor	French Suite No. 2, BWV 813
5		D	Prelude No. 5, BWV 850 (Well-tempered Piano I)
6		F	Little Fugue, BWV 556
7		G	French Suite No. 5, BWV 816
8	Mozart	A	Sonata K331, Allegro
9		B ^b	Sonata K281, Allegro
10		C	Sonata K545, Allegro
11		C-minor	Sonata K457, Allegro assai
12		D	Sonata K576, Allegro
13		F	Sonata K280, Allegro
14		G	Sonata K283, Allegro
15	Beethoven	A	Sonata No. 2, Op. 2, Allegro
16		B ^b	Sonata No. 11, Op. 22
17		C	Sonata in C, Op. 21, Allegro con brio
18		C-minor	Sonata No. 5, Op. 10 No. 1, Allegro
19		D	Sonata No. 7, Op. 10, Presto

20	F	Sonata No. 6, Op. 10 No. 2
21	G	Sonata No. 10, Op. 14, Allegro

891 *Note.* All are major keys except those explicitly designated as minor.

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

Appendix B

913

Experiment 2: Natural Stimuli

No.	Composer	Pianist	Key	Title
1	Bach	Arrau		Partita No. 2: Rondeaux Philips 434 904-2
2				Partita No. 3: Fantasia Philips 434 904-2
3				Partita No. 5: Praeambulum Philips 434 904-2
4		Barenboim		Goldberg Variations: Var. 18 Erato 741397T
5				Goldberg Variations: Var. 5 Erato 741397T
6				Goldberg Variations: Var. 6 Erato 741397
7		Pirès		Partita No. 1: Praeludium Philips 456 928-2
8				English Suite No. 3: Prelude Philips 456 928-2
9				French Suite No. 2: Allemande Philips 456 928-2
10		Richter		Das Wohltemperierte Clavier, Book I: Prelude 2 RCA Victor GD 60949

11			Das wohltemperierte Clavier, Book I: Prelude 5 RCA Victor GD 60949
12			Das wohltemperierte Clavier, Book II: Prelude 6 RCA Victor GD 60949
13	Mozart	Arrau	Sonata, KV 284: mvmt 1 Philips 432 306-2
14			Sonata, KV 330: mvmt 1 Philips 432 306-2
15			Sonata, KV 576: mvmt 1 Philips 432 306-2
16		Barenboim	Sonata, KV 281: mvmt 1 EMI CDZE 7 67294 2
17			Sonata, KV 533: mvmt 1 EMI CDZE 7 67294 2
18			Sonata, KV 311: mvmt 1 EMI CDZE 7 67294 2
19		Pirès	Sonata, KV 280: mvmt 1 DG 435 882-2
20			Sonata, KV 282: mvmt 3 DG 435 882-2
21			Sonata, KV 333: mvmt 1 DG 435 882-2
22		Richter	Sonata, KV 283: mvmt 1

			Philips 438 480-2
23			Sonata, KV 310: mvmt 1
			Philips 422 583-2
24			Sonata, KV 457: mvmt 1
			Philips 438 480-2
25	Beethoven	Arrau	Sonata No. 15, op. 28: mvmt 1
			Philips 426 068-2
26			Sonata No. 21, op. 53: mvmt 1
			Philips 426 068-2
27			Sonata No. 26, op. 81a: mvmt 3
			Philips 426 068-2
28		Barenboim	Sonata No. 22, op. 54: mvmt 2
			EMI 5 72912 2
29			Sonata No. 11, op. 22: mvmt 1
			EMI 5 72912 2
30			Sonata No. 28, op. 101: mvmt 2
			EMI 5 72912 2
31		Pirès	Sonata No. 14, op. 27, no. 2: mvmt 3
			Erato 3984 27487 2
32			Sonata No. 17, op. 31, no. 2: mvmt 3
			Erato 3984 27487 2
33			Sonata No. 23, op. 57: mvmt 3
			Erato 3984 27487 2

34	Richter	Sonata No. 7, op. 10, no. 3: mvmt 1 Praga 354 022
35		Sonata No. 3, op. 2, no. 3: mvmt 1 Brilliant 92229/3
36		Sonata No. 31, op. 110: mvmt 2 Philips 454 949-2

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931 Table 1.

932 *Variance (%) explained by the first four components in Experiments 1 and 2 in terms of*

933 *composers.*

934

	Experiment 1		Experiment 2	
	Free	Constrained	Free	Constrained
Component-1	13.83	19.70	10.58	21.47
Component-2	11.03	13.21	7.93	13.03
Component-3	8.06	9.44	5.89	7.03
Component-4	6.82	7.83	5.28	5.76

935

936

937

938

939

940

941

942

943

944

945 Figure Captions

946 *Figure 1.* Compromise factor scores for Experiment 1: Free sorting task with MIDI stimuli,
947 color-coded by composers (Bach: purple; Mozart: green; Beethoven: lavender). Panel (a)
948 Components 1 and 2, and Panel (b) Components 3 and 4.

949 *Figure 2.* R_V factor scores for Experiment 1: Free sorting task with MIDI stimuli, color-coded by
950 musical experience (highly trained: orange; moderately trained: red; untrained: blue). Each dot
951 represents a participant and the numbers corresponding to each dot represent the years of music
952 training.

953 *Figure 3.* Compromise factor scores for Experiment 1: Constrained sorting task with MIDI
954 stimuli, color-coded by composers (Bach: purple; Mozart: green; Beethoven: lavender). Panel (a)
955 Components 1 and 2, and Panel (b) Components 3 and 4.

956 *Figure 4.* R_V factor scores for Experiment 1: Constrained sorting task with MIDI stimuli, color-
957 coded by musical experience (highly trained: orange; moderately trained: red; untrained: blue).
958 Each dot represents a participant and the numbers corresponding to each dot represent the years
959 of music training.

960 *Figure 5.* Compromise factor scores for Experiment 2: Free sorting task with natural stimuli,
961 color-coded by composers (Bach: purple; Mozart: green; Beethoven: lavender). Panel (a)
962 Components 1 and 2, and Panel (b) Components 3 and 4.

963 *Figure 6.* Compromise factor scores for Experiment 2: Free sorting task with natural stimuli,
964 color-coded by pianists (Arrau: purple; Barenboim: orange; Pirès: pink; Richter: red). Panel (a)
965 Components 1 and 2, and Panel (b) Components 3 and 4.

966 *Figure 7.* R_V factor scores for Experiment 2: Free sorting task with natural stimuli, color-coded
967 by musical experience (highly trained: orange; moderately trained: red; untrained: blue). Each

968 dot represents a participant and the numbers corresponding to each dot represent the years of
969 music training.

970 *Figure 8.* Compromise factor scores for Experiment 2: Constrained sorting task with natural
971 stimuli, color-coded by composers (Bach: purple; Mozart: green; Beethoven: lavender). Panel (a)
972 Components 1 and 2, and Panel (b) Components 3 and 4.

973 *Figure 9.* Compromise factor scores for Experiment 2: Constrained sorting task with natural
974 stimuli, color-coded by pianists (Arrau: purple; Barenboim: orange; Pirès: pink; Richter: red).
975 Panel (a) Components 1 and 2, and Panel (b) Components 3 and 4.

976 *Figure 10.* R_V factor scores for Experiment 2: Constrained sorting task with natural stimuli,
977 color-coded by musical experience (highly trained: orange; moderately trained: red; untrained:
978 blue). Each dot represents a participant and the numbers corresponding to each dot represent the
979 years of music training.

980

981

982

983

984

985

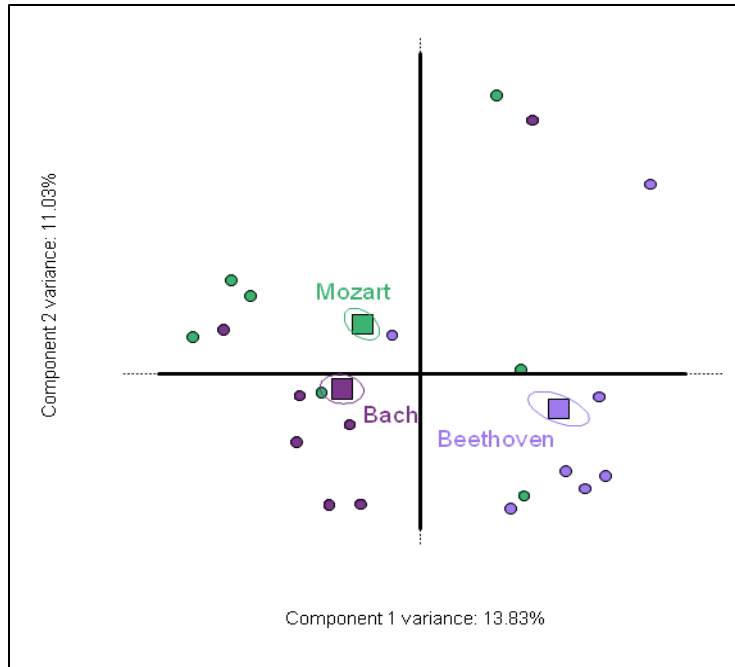
986

987

988

989

990

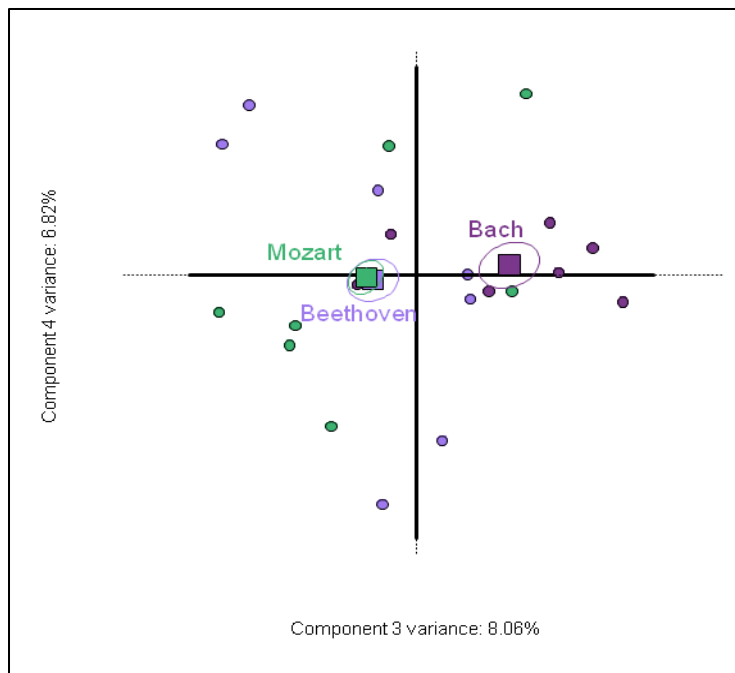


991

992

993

(a)



994

995

996

(b)

Figure 1.

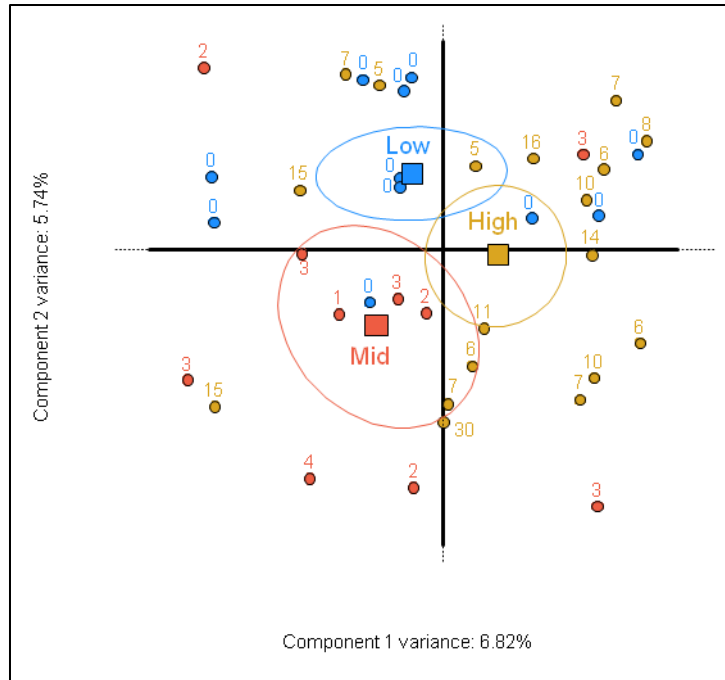


Figure 2.

997

998

999

1000

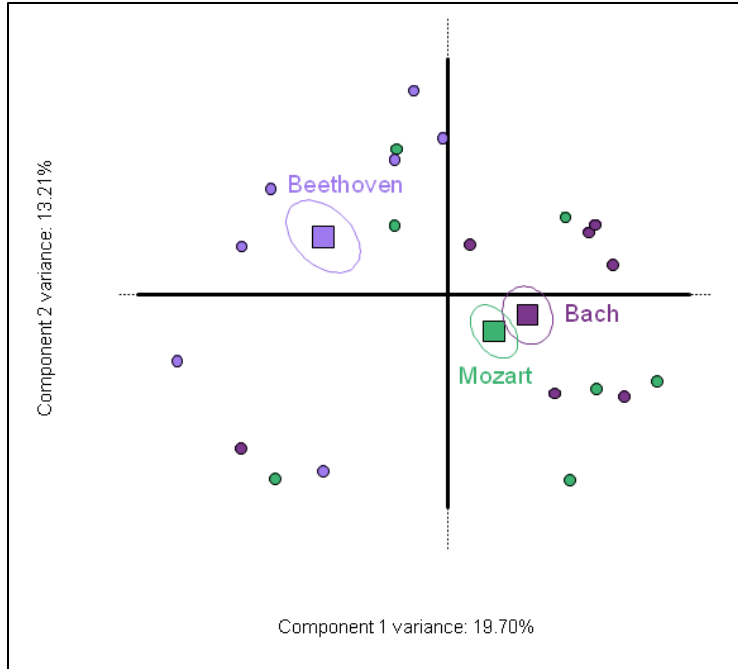
1001

1002

1003

1004

1005

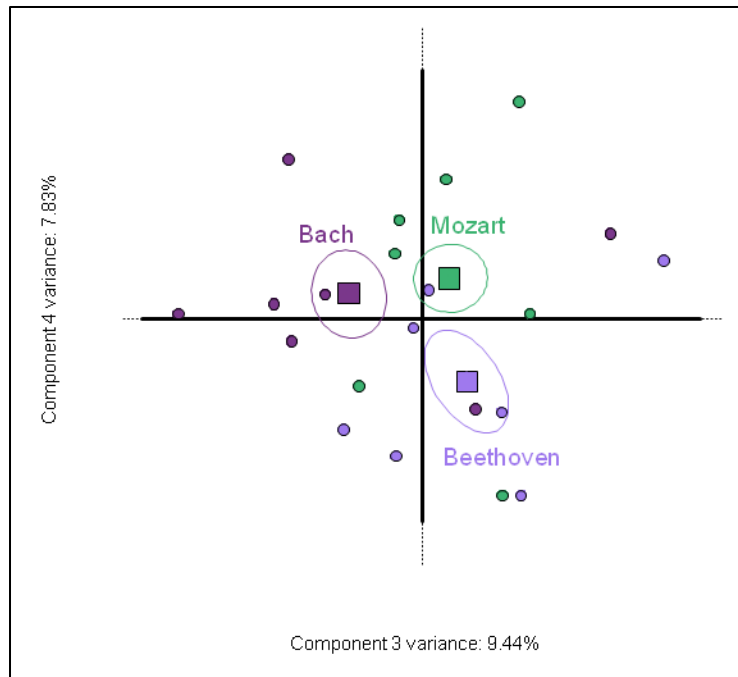


1006

1007

1008

(a)



1009

1010

1011

(b)

Figure 3.

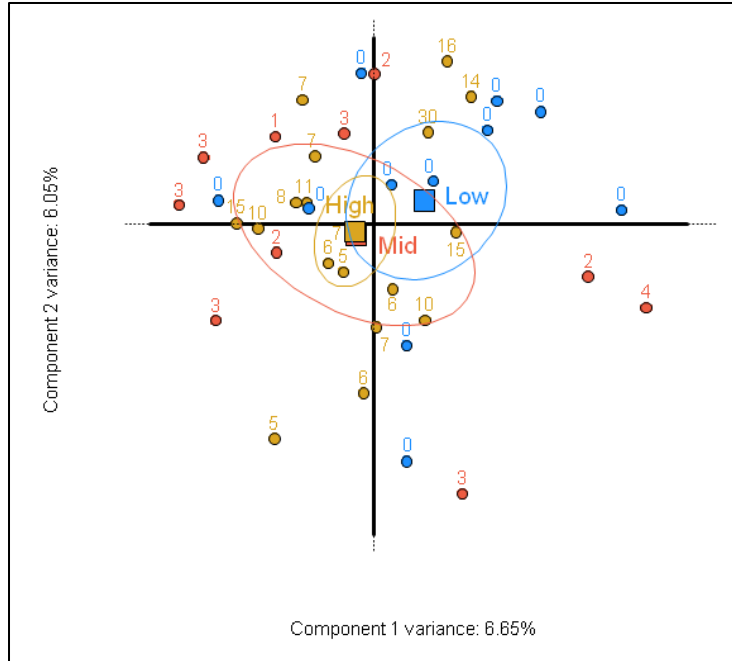


Figure 4.

1012

1013

1014

1015

1016

1017

1018

1019

1020

1021

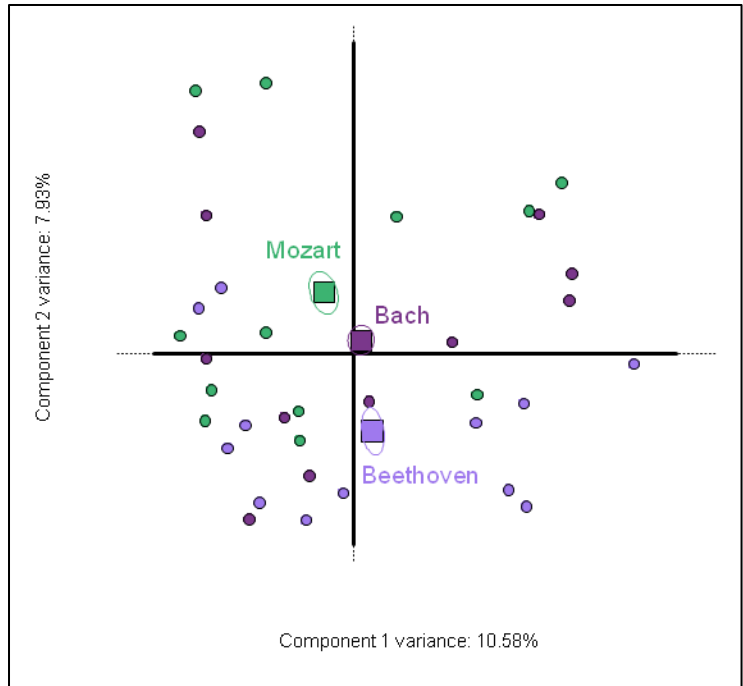
1022

1023

1024

1025

1026

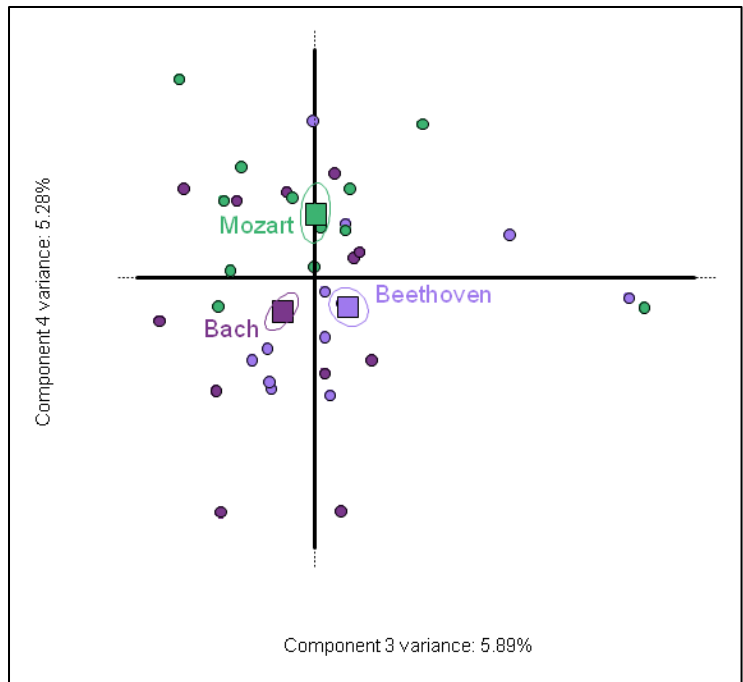


1027

1028

1029

(a)



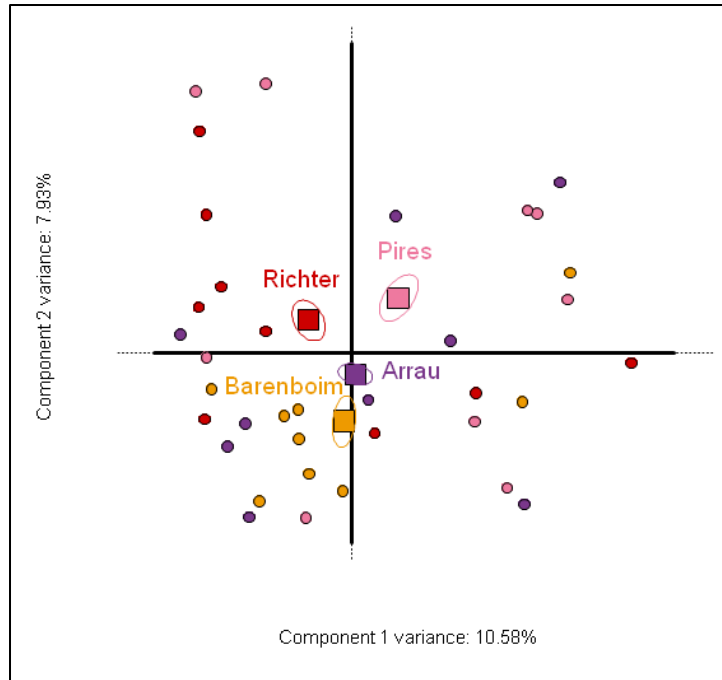
1030

1031

1032

(b)

Figure 5.

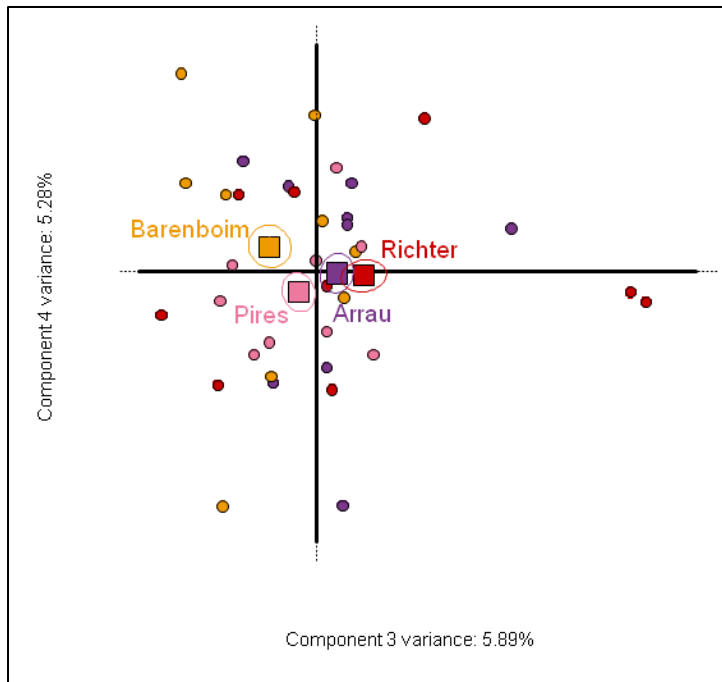


1033

1034

1035

(a)



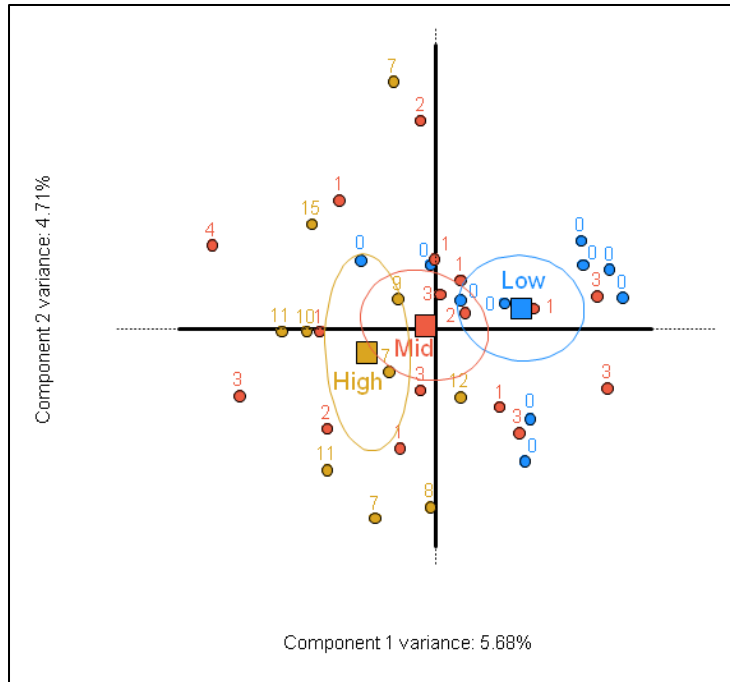
1036

1037

1038

(b)

Figure 6.



1039

1040

1041

1042

1043

1044

1045

1046

1047

1048

1049

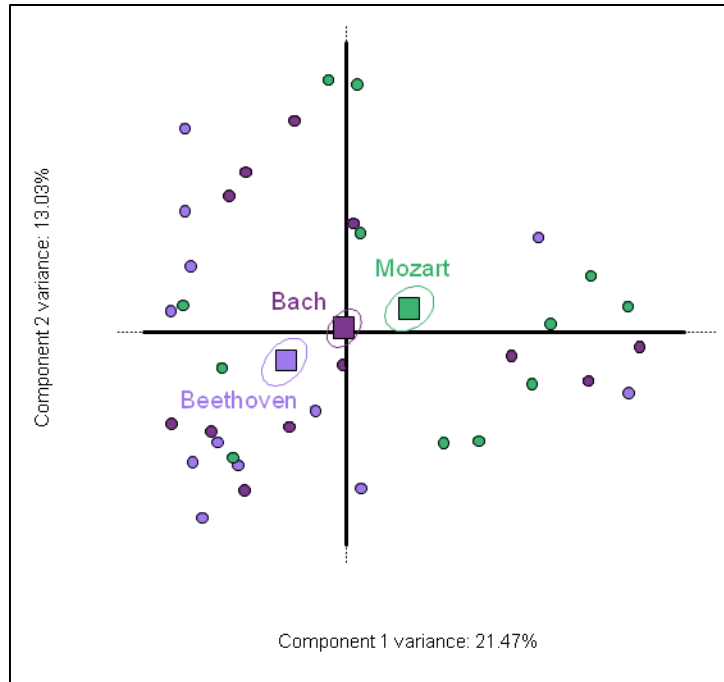
1050

1051

1052

1053

Figure 7.

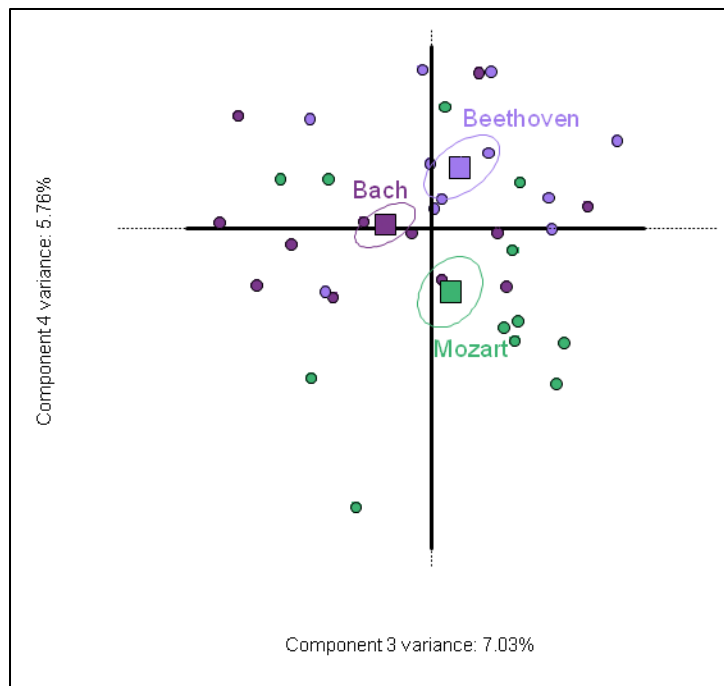


1054

1055

1056

(a)



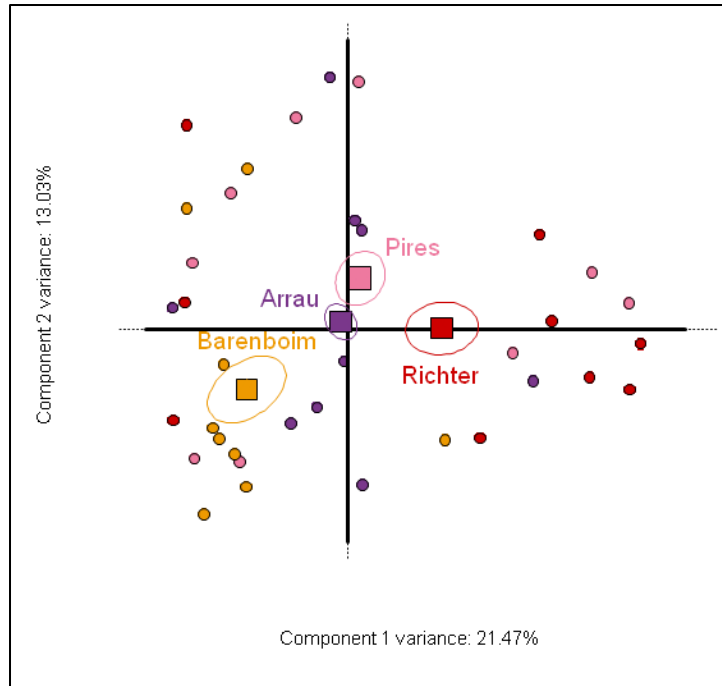
1057

1058

1059

(b)

Figure 8.

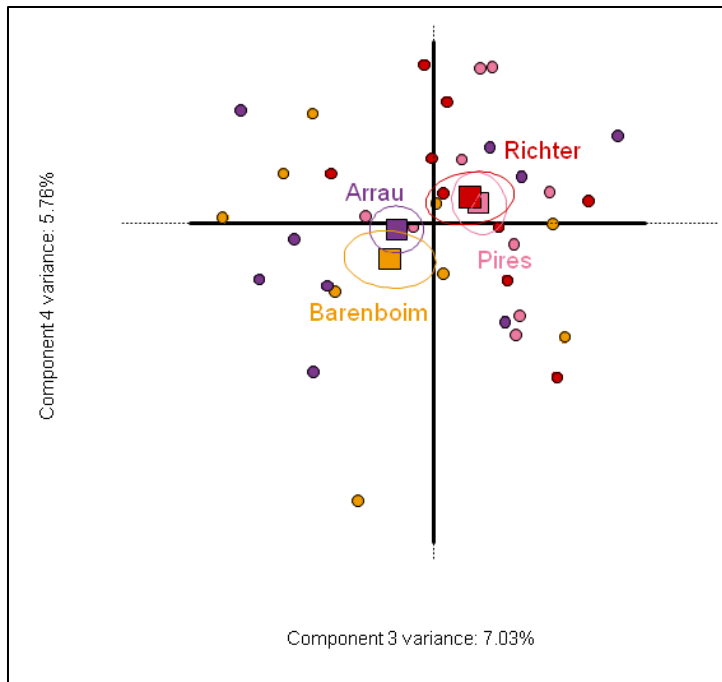


1060

1061

1062

(a)



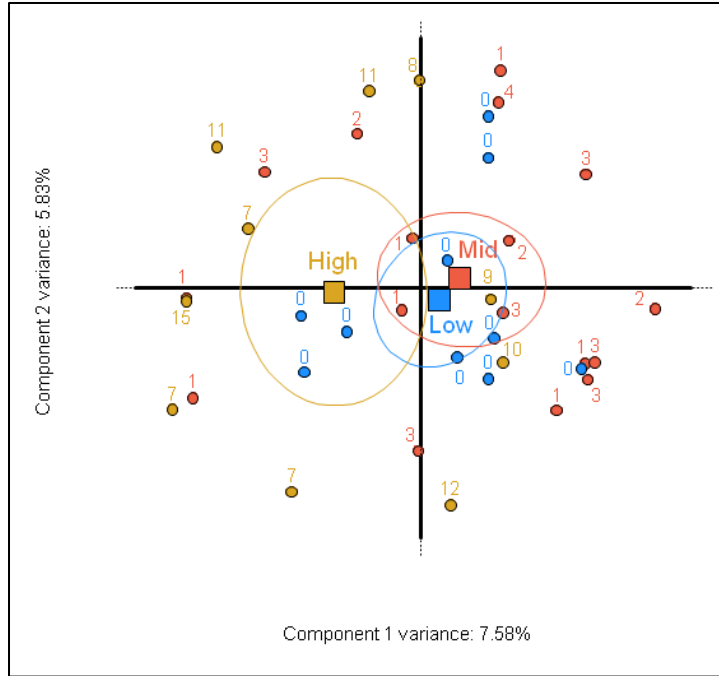
1063

1064

1065

(b)

Figure 9.



1066

1067

1068

1069

1070

1071

Figure 10.