"It is Different Each Time I Play": Variability in Highly Prepared Musical Performance

Roger Chaffin
University of Connecticut

Anthony F. Lemieux
Purchase College, State University of New York

Colleen Chen
University of Connecticut

PERFORMANCES BY CONCERT PIANISTS in the Western classical tradition are normally highly prepared, yet must sound fresh and spontaneous. We propose that musicians achieve the necessary spontaneity by strategic management of the variability inherent in any action. Musical gestures that make up the artist's interpretation (e.g., crescendos, ritardandos) are attenuated or exaggerated to different degrees in each performance, while movements critical for technique are less varied. We examined 7 highly polished performances of J.S. Bach's *Italian Concerto (Presto)* by a concert pianist. There were small but consistent differences between performances in 4 of 9 identified musical gestures, each of which occurred in several locations. In contrast, at points where the pianist reported attending to technique during performance, slower tempi and lower dynamic variability suggested that she controlled execution of planned movements more closely. Increased control at technical difficulties permitted more spontaneous variation in the musical gestures important to her interpretation.

Received November 16, 2005, accepted December 13, 2006.

Key words: skilled performance, musical interpretation, motor control, musical creativity, performance cues

"IT IS DIFFERENT EACH TIME I PLAY," reported the great Russian pianist Emil Gilels (Mach, 1991, p. 123). The impression that repeated performances of the same piece by the same artist differ in musically significant ways is widespread and perhaps contributes to the continuing popularity of live performance (Ross, 2005). Is the impression correct, and if so, are the differences due to random variation or to something more systematic that is governed by musical intuition? On the one hand, musical performance is subject to the same kind of random variability as any other psychological process (Bernstein, 1967; Gilden, 2001; Newell & Corcos, 1993). Perhaps performers and audiences alike simply ignore or forget most of the differences between performances and remember only those random fluctuations that happen to be musically relevant. On this view, there are differences but they are just a subset of the random fluctuations that permeate any motor activity. Alternatively, some of the differences may be a product of the performer's musical spontaneity and creativity. The larger context of Gilels' remark makes it clear that he takes the latter view, believing that the changes he describes reflect artistry and skill:

> "When I am in top form . . . the ideas are always different. Sometimes I play with greater changes in dynamics, sometimes with less . . . I must say it is different each time I play, and it is a process which . . . includes mastery of the work, knowing the details, being comfortable with it, and then adding the fantasy (Emil Gilels, in Mach, 1991, p. 123).

Our goal was to understand the basis for the intuition that Gilels articulated by examining repeated performances of a highly prepared piece in which the performer was giving the same interpretation each time. Artists are able to reliably reproduce the same nuances of interpretation in repeated performances and normally do so, once a new piece has been prepared, giving essentially the same performance on more than one occasion (Clynes & Walker, 1982, 1986; Shaffer, 1984; Shaffer, Clarke & Todd, 1985; also see Gabrielsson, 1999, for a review). We were interested in differences that occur as a normal part of repeating the same performance, not in the kind of deliberate changes that a performer might make when exploring the interpretive possibilities of a piece (Clarke, 1995) or when relearning a piece with the goal of rethinking its interpretation (Repp, 1992). We hypothesized that, if Gilels is right and there are spontaneous differences between repeated performances, they

Music Perception VOLUME 24, ISSUE 5, PP. 455–472, ISSN 0730-7829, ELECTRONIC ISSN 1533-8312 DOI:10.1525/MP.2007.24.5.455

are a product of the normal motor and cognitive processes responsible for performance. In this case, such differences should be present in any set of repeated performances, whether the musician intends the changes or not, and whether the musician is aware of them or not. We will ask if such differences are simply random or whether they are a reflection of musical spontaneity, or, as Gilels put it, of the artist's musical mastery and musical imagination.

What differences between repeated performances might reflect musical spontaneity as opposed to random variability? Gilels gives an example of what to look for when he says, "Sometimes I play with greater changes in dynamics, sometimes with less . . ." (Mach, 1991, p. 123). He apparently believes that his playing on different occasions differs in systematic and global ways, such as in the magnitude of dynamic changes. If this is true of performances in general, then we should be able to find global properties of this sort that differ systematically from one performance to another. Alternatively, if differences between repeated performances are random, there will be no systematic differences. Instead, properties like degree of dynamic change will vary randomly within each performance, larger than normal in some places and smaller than normal in others. Differences between performances will simply be part of this random variation. Because statistically reliable differences between performances must be identified relative to variability within performances, random variation will produce no systematic differences of the sort that Gilels describes.

We set ourselves the challenge of finding systematic differences between performances that were as similar as possible in order to avoid any changes in interpretation that a musician might deliberately decide to make from one performance to the next. If systematic differences between performances do occur, they should be present in any set of repeated performances, whether the musician is aware of them or not. We examined seven performances of the third movement (*Presto*) of J.S. Bach's *Italian Concerto* by pianist Gabriela Imreh recorded as part of a larger study (Chaffin & Imreh, 1997, 2001, 2002; Chaffin, Imreh & Crawford, 2002; Chaffin, Imreh, Lemieux, & Chen, 2003; Imreh & Chaffin, 1996/97; see Chaffin & Logan, 2006, for a review). The performances took place near the end of a 10-month period during which the pianist prepared the piece for a professional recording (Imreh, 1996), after learning of the piece was complete. The performances all took place in the pianist's practice studio, without an audience, as part of the pianist's normal preparation. The goal in each case was to play as well as possible. The pianist judged the performances to be very similar to each other and close to her ideal performance and reported that they differed only in the occasional wrong note and how "cautious" they sounded. Thus they differed only in ways that would be found with any repeated performances of the same piece over a short time-span.

We interpret Gilels' remark as suggesting that when the same piece is repeated, the repetitions will often differ in the degree to which the various musical gestures of the interpretation are emphasized. Musical interpretation consists of deviations from the strict regularity or equality in timing, dynamics, and articulation that a literal rendition of a score would provide. Musical gestures are the patterns of deviation that embody each performer's unique interpretation (e.g., crescendos, ritardandos, micro-pauses between phrases) and are reliably reproduced from one performance to the next (Clynes & Walker, 1982, 1986; Shaffer et al., 1985; see Clarke, 2004, for a review). We speculate that variation in musical gestures, both within and across performances, is one aspect of keeping a performance fresh and spontaneous. When a gesture is repeated at different points in the same performance, varying its emphasis adds novelty and interest for the listener. When a performance is repeated, the same kind of change from one performance to another may provide novelty and interest for the musician.

We empirically identified musical gestures in the performances under study by noting correspondences between musical features in the score and fluctuations in tempo and loudness in the performances. For example, the main theme of the *Presto* begins with a downward octave jump that was marked by a sharp decrease in tempo (see Figure 1). This musical gesture occurred six times, at different points throughout the piece. A difference in this gesture between performances would thus be the kind of systematic, global difference that we are looking for. The difference will not be easily attributed to random changes in a single bar but will indicate that the gesture was systematically exaggerated or diminished at several places in the performance. All of the musical gestures involved features of this sort that occurred at several points in the piece. Differences between performances would thus represent the kind of difference that Gilels believed to be characteristic of his performances.

Our method was correlational; therefore, it is possible that any effects we obtain may be due to properties of the music other than the ones we have identified. To test the plausibility of alternative explanations we will briefly report the results of additional

FIGURE 1. Score of bars 1-3 showing the octave jump down (Jump: *A*) in bar 1 and the ascending 8th note scale in bar 2.

analyses including predictors for every aspect of technique, interpretation, performance, and musical structure that the pianist considered relevant to her learning and performance of the piece. The pianist had provided these reports as part of the larger study of the *Presto* (Chaffin et al., 2002). Including all the musical properties that the pianist was able to identify decreases the probability that any effects we do obtain are due to other, unidentified musical properties.

Trade-offs Between Flexibility and Stability

Repetitions of the same movement are never identical but are always performed differently (Bernstein, 1967). This variability arises from the fact that the motor system has multiple ways to execute any movement. This redundancy is responsible for both the stability of the motor system and its flexibility. Errors in one part of the system are corrected by adjustments in other parts to achieve the intended movement. Trade-offs between stability and flexibility are thus a fundamental property of the motor system (Latash, Scholz, & Schöner, 2002; Todorov & Jordan, 2002). Expert performers in a variety of domains appear to capitalize on this property of the motor system and engage in strategic trade-offs between flexibility in some aspects of performance in return for increased stability in others. For example, long-jumpers vary their stride as they approach the jump-off plate in order to achieve precision on the final take-off (Lee, Lishman, & Thomson, 1982), and marksmen control the position of the gun on dimensions that determine bullet trajectory (pitch and yaw) more closely than on dimensions that do not (roll and position along the line of shooting; Scholz, Schöner, & Latash, 2000). If musicians also engage in strategic trade-offs between flexibility and stability, then we would expect that the flexibility of musical gestures that we have proposed would be complemented by increased stability in other aspects of musical performance.

Which aspects of a polished performance would we expect to be sources of stability? One possibility is the technically difficult passages that a musician plays repeatedly during practice in order to develop the necessary motor skills, e.g., a long jump across the keyboard. Such passages might become more stereotyped and thus more stable than others requiring less practice. Another possibility is the *basic performance cues* that a musician attends to during performance in order to ensure that details of technique are executed as planned, e.g., a fingering necessary to position the hand for what is to follow (Chaffin et al., 2002, pp. 169-172). Learning a new piece involves making myriad detailed decisions about many aspects of technique, such as fingerings. Most of these become automatic through practice. Experienced performers, however, single out a small subset of these decisions where doing things as planned rather than in some other way is crucial for what follows. By learning to pay attention to cues during performance, the performer makes sure that they are executed precisely. These cues, together with others representing the expressive and interpretive turning points of the piece, provide the musician with a mental map that is used to track progress as the performance unfolds. Basic performance cues may be sources of stability in a performance because they ensure that particular actions are done the same way in each performance.

As part of the study of the *Presto* (Chaffin et al., 2002), the pianist provided detailed reports about every aspect of the music that she thought about during practice, including the location of basic performance cues and technical difficulties. Analyses of her practice showed that she started at, stopped at, and repeated passages containing basic cues and technical difficulties more than other places (Chaffin et al., 2002, pp. 179-190). Here we ask whether the polished performances were more stable at these points. If technical difficulties are more stable, this will suggest that the increased attention during practice was responsible. If basic performance cues are more stable, this will suggest that increased attention during performance was responsible.

We expected that increased stability in technique would be distinguished by a decrease in *dynamic variability* because the pianist would have fewer attentional resources to devote to subtle interpretive effects in passages where she needed to pay attention to technique. The *Presto* is typical of the polyphonic music of the Baroque era in possessing a dense sound texture in which several voices run in parallel. To provide a rewarding experience for the listener, a performer must give each voice a distinct character (e.g., staccato or legato). Playing all the notes of a particular voice in the same way unites them in the listener's experience, making them stand out as a continuing pattern. A large part of the performer's art is to draw the listener's attention from one voice to another in turn by varying the emphasis on each. We hypothesized that the pianist's ability to make the subtle adjustments needed to do this would decrease at sources of stability, and that the range of dynamic levels would be momentarily reduced as a result. In addition, it was possible that tempo would decrease in these passages in order to allow more time to ensure the accurate execution of these critical passages.

In summary, we looked for both flexibility and stability in repeated performances that the musician felt were very similar. Finding reliable differences would suggest that similar effects would be present in any set of repeated performances and document the intuition expressed by Gilels that every performance is different. Finding sources of stability would represent a possible first step toward providing a psychological explanation for the differences.

Method

Learning the Presto

The pianist videotaped her practice as she learned the *Italian Concerto* for the first time for a professional recording. The *Presto* is of moderate difficulty (Hinson, 1987), is scored in 210 bars divided into 16 main sections and 37 sub-sections, is notated in 2/2 time, and lasts for 3-4 minutes at performance tempo. The score is available online (http://www.music-scores.com/ and http:// www.musicnotes.com/) and in Chaffin et al., (2002, pp. 272-278).

The Performances

We selected seven uninterrupted performances from near the end of the 33 hours, 57 sessions, and 42 weeks of practice: one from Session 42 (42:2), five from Session 49 (49:2, 49:3, 49:4, and 49:5), and two from Session 50 (50:1 and 50:2). The performance from Session 42 (42:2) was included because the pianist had commented at the end of the previous session, "There isn't that much more that I can do," indicating that the development of interpretation was complete, and at the end of Performance 42:2 remarked, "It's getting better," suggesting that the performance had been satisfactory. The pianist stopped recording her practice after Session 45 because she felt that preparation of the piece was complete, but continued to play through the piece a few times each day in order to maintain it in readiness for the recording session (Chaffin et al., pp. 113-114). In Sessions 49 and 50 she taped two of these sessions in order to make a recording for use during talks about the research. The pianist considered all of the performances to be close to her ideal although none matched it exactly, either because of small mistakes or because they were "*overly cautious*."

Tempo, Mean Dynamic Level, and Dynamic Variability

Inter-bar intervals (IBIs) were measured from the start of one bar to the start of the next using a commercial sound wave processing program applied to the sound wave of each performance. Measurements that differed from the mean performance by more than two standard deviations were repeated. Reliability of the resulting measurements was assessed by the same person making independent measurements of passages from two of the performances totaling 70 bars. The mean absolute errors of measurement were 11 and 14 ms ($SD = 10$ and 11 respectively). Table 1 shows the distribution of measurement errors represented by the differences between the two sets of measurements. More than 90% of differences were less than 29 ms or 3.3% of the mean IBI.

For the following analyses, IBIs were converted to tempo measured in beats per minute (tempo = (1/IBI in seconds) × 2 beats per bar × 60 s/min).

Dynamic levels were measured by dividing the sound signal into 10 ms slices and computing perceived intensity levels for each slice in sones (Zwicker & Fastl, 1999). IBI measurements were then used to divide the series of sound slices into bars, with a mean of 84 slices per bar.

TABLE 1. Summary of Inter-Bar Interval (IBI) reliability measurements showing cumulative percentage at each range of difference.

	Performance	
Range of difference (ms)	49:2	49:5
0-5	32.3	30.0
6-10	58.5	48.6
11-15	76.9	64.3
16-20	84.6	77.1
21-25	89.2	80.0
26-30	90.8	92.9
31-35	98.5	98.6
36-40	—	—
41-45	100.0	—
46-50	—	—
51-55	—	—
56-60	—	100.0

The mean and standard deviation of these values provided the measures of mean *dynamic level* (in sones) and *dynamic variability* (standard deviation of dynamic level) for each bar. Mean tempo, dynamic level, and dynamic variability were computed across the seven performances for each bar. The resulting mean profiles were first used to edit the performances, as described below, and then recomputed for analysis.

There were seven places where the pianist omitted notes, played a wrong note, or noticeably hesitated, one in each performance except 49:5 and two in 49:2. Data for these bars, and one or two bars following were replaced with the corresponding values from the appropriate mean profile normalized to the mean of the edited performance. In all, 20 bars were edited, 1.2% of the total. The procedure was conservative with respect to the hypothesis of differences between performances since it reduced differences.

Identification of Musical Gestures

We identified musical gestures as deviations in tempo or mean dynamic level that were systematically related to musical features in the score throughout the piece and across the seven performances. The musical features were identified in two different ways. Four were identified directly by the investigators by finding fluctuations in mean tempo or dynamic level that corresponded to musical features, e.g., the downward octave jump in the *A* theme mentioned in the introduction. Five were identified using reports that the pianist provided as part of an earlier study as a guide to where to look for tempo and dynamic level fluctuations. In each case, the musical gestures were identified empirically by preliminary analyses of the type described in detail below. Sources of stability were identified at a later stage of the investigation, after the musical gestures had been identified, by looking for correspondences between fluctuations in dynamic variability and the pianist's reports of decisions involving technique.

As part of an earlier study, the pianist provided detailed reports of her musical understanding of the piece by marking features of the music on copies of the score for 12 separate dimensions dealing with musical structure, performance, interpretation, and technique, (see Chaffin et al., 2002, pp. 171-176 for details). The dimensions reported by the pianist that were used in the present study were: boundaries between sections containing different musical themes, basic and expressive performance cues, beginnings of phrases, and technical difficulties. These reports were used as a source of hypotheses about the location of musical gestures and sources of stability. For example, dynamic variability increased in bars where the pianist reported beginnings of phrases; phrasing was therefore included as a musical gesture.

The nine musical gestures identified are listed in Table 2 along with two potential sources of stability. The table shows for each gesture and source of stability whether the musical feature involved was reported by the pianist or identified by the investigators. The table also shows the range of values used to code each bar for each of the musical features and the number of bars with non-zero codings. All of the gestures occurred at more than one location and three gestures encompassed the

TABLE 2. Descriptions of musical gestures and potential sources of stability.

	Description	Source	Values coded	Number of bars coded > 0
Musical gestures				
Jump: *A*	Octave jump down at start of *A* theme	Investigator	0,1	6
Other jumps	Other large downward jumps	Investigator	0,1	5
Scales	Octave, 8th notes before or after Jump: *A*	Investigator	0,1	15
Sections	Sections of main themes: *A,B,C* or *D*	Pianist	1-4	210
Section ends	Last bar of subsections	Pianist	0,1	37
Position in section	Serial position from start of subsection	Pianist	1-19	210
Modulation	Change of key	Investigator	0,1	10
Phrasing	Starts of phrases	Pianist	0-7	188
Expressive intensity	Expressive intensity reported by pianist	Pianist	1-5	210
Sources of stability				
Basic performance cues	Attention to technique in performance	Pianist	0-3	125
Technical difficulties	Technique required repetition during practice	Pianist	0,1	126

entire piece.[1] The musical gestures identified by the investigators were all represented by dummy variables, as were the technical difficulties and ends of sections reported by the pianist.[2] The four themes were coded by ranking them on mean tempo and then assigning each bar the rank associated with the theme of which it was a part.[3] Position of each bar in a section was coded as serial position from the beginning of the subsection. Basic performance cues were coded by the number of cues reported for each bar, and phrasing was coded as the number of phrases starting in each bar.[4] The pianist provided ratings of *expressive intensity* using a 5-point scale with 5 representing high intensity. As part of the earlier study, the pianist had reported the location of expressive performance cues, thus identifying places where her expressive goals changed, e.g., from “light but mysterious” to “surprise” (Chaffin et al., 2002, p. 172). For the present study, she provided additional information about each expressive cue by rating the overall emotional intensity of the passage that it introduced (Krumhansl, 1997). The rating was assigned to the bar containing the cue and to each following bar until the next expressive cue.

Analysis

Multiple regression analyses were performed in which the predictor variables represented musical gestures and sources of stability (see Table 2) and the predicted variables were tempo, mean dynamic level, and dynamic variation. The bars were the unit of analysis ($N = 210$). Each bar was coded for the presence of sources of stability (technical difficulties and basic performance cues) and nine musical gestures.

Three sets of analyses were performed. First, the effects of musical gestures and potential sources of stability were examined for the mean profiles for tempo, dynamic level, and dynamic variability. Second, the individual performances were compared with the mean performance. Differences were identified by entering the mean performance as an additional predictor along with the predictors representing gestures and sources of stability. Significant effects would represent differences of the individual performance from the mean for a particular gesture or source of stability (see Campbell & Kenny, 1999, for discussion of techniques for assessing differences). Alternatively, if the only differences between performances were due to random noise, there would be no significant effects for these predictors. Third, the analysis in stage 2 was performed again in step-wise fashion with the mean performance, sources of stability, and musical gestures entered in three steps to determine how much of the variability in each performance was attributable to each source.

For the first two sets of analyses, supplementary analyses tested the possibility that effects might be due to other, uncontrolled variables not included in the main analyses by adding predictors for additional musical properties quantified in earlier studies of the *Presto* (number of notes, and reports by the pianist of fingerings, familiar patterns, dynamic changes, tempo changes, pedaling, and interpretive performance cues; Chaffin et al., 2002, Ch. 8).

Results & Discussion

Descriptive Statistics for the Performances

As required by the score and by Baroque performance practice, the tempo was fast and variation in tempo and dynamic level was minimal (see Table 3). Mean tempo within performances ranged from almost 140 to 143 beats/min with standard deviations of 4-5 beats/min, approximately 3% of mean tempo. Mean dynamic levels within performances ranged from 39.6 to 45.6 sones, and dynamic variability ranged from 6.2 to 7.4 (standard deviations). Although the values were very similar across performances, the differences between performances were consistent, $F(6,1236) = 29.68$, $p < .001$, $F(6,1242) = 263.58$, $p < .001$, and $F(6,1242) = 44.87$, $p < .001$ for tempo, mean dynamic level, and dynamic variability respectively.

The tempo and dynamic profiles of the seven performances were generally similar (see Figures 2 and 3). For example, every performance in Figure 2 shows a downward spike (slowing) in bar 12 and a higher (faster) section starting in bar 77. There were also differences. For example, the downward spike in bar 12 was much reduced in performance 42:2; the higher

[1] The number of bars coded with non-zero values ranged from 5 to 210, showing that each gesture and potential source of stability involved multiple bars. Gestures coded with continuous values (serial position and expressive intensity) had non-zero values in all 210 bars, as did the ranking of the four themes in terms of their mean tempo.

[2] The pianist’s report of the formal structure divided the piece into major sections (see footnote 3), which were further subdivided into 37 subsections, which we call “sections” here for ease of reference.

[3] The four themes were ranked from fast to slow: C, B, A, D. The following major sections were identified in terms of the four themes as follows: A (1-12), A1 (13-24), B (25-44), B1 (45-64), A2 (65-76), C (77-92), A3 (93-103), D (104-122), A4 (123-126), B2 (127-138), A5 (139-149), A6 (150-154), C1 (155-166), B3 (167-186), A (187-198), A1 (199-210).

[4] The pianist marked phrases separately for each voice and so each bar typically contained starts of more than one phrase.

TABLE 3. Tempo, mean dynamic level, and dynamic variability for seven performances.

	Tempo (beats/min)*		Mean dynamic level (sones)**		Mean dynamic variability (sones)**	
Performance	*M*	*SD*	*M*	*SD*	*M*	*SD*
42:2	141.31	4.97	43.93	4.32	7.07	1.44
49:2	143.05	5.25	39.60	4.23	6.24	1.38
49:3	143.10	5.49	42.21	4.64	6.85	1.51
49:4	142.08	5.17	41.89	4.27	6.90	1.53
49:5	141.71	5.17	42.87	4.83	6.60	1.47
50:2	139.89	5.39	44.44	4.77	7.06	1.66
50:3	140.44	5.56	45.65	4.76	7.39	1.71
M	141.66	4.37	42.94	4.25	6.87	1.34

*last three bars omitted, **first and last bar omitted

region starting at bar 77 appears more pronounced in 49:2 and barely present in 50:3. The correlations between the profiles ranged from $r(207) = .54$ to .71, $p <$ .001, for tempo and from $r(208) = .77$ to .91, $p < .001$, for dynamic level and variability. Inclusion of the final three bars for tempo and the first and last bars for dynamic level increased most of these values to $r > .90$, masking differences between the profiles. Therefore, these bars were omitted from Figures 2 and 3 and from all analyses.

Effects of Musical Gestures and Sources of Stability on the Mean Performance

To test whether the musical gestures and sources of stability reflected statistically reliable correlations between features of the score and fluctuations in the temporal and dynamic profiles, we used multiple regression analyses to assess the effects of the musical gestures and sources of stability on tempo, mean dynamic level, and dynamic variability. Table 4 summarizes the analyses and shows the number of significant effects for the individual performances (all in the same direction as for the means). The musical gestures affected the mean performance on at least one dependent measure except for serial position in a section and modulation, which only affected tempo for one and two of the individual performances respectively. In all, we were able to account for 46% of the variance in tempo, 39% in mean dynamic level, and 23% in dynamic variability. The low autocorrelations indicated that the data met the assumption of independence required for the statistical tests reported below; adjacent bars were unrelated for tempo and dynamic variability and the correlation for dynamic level was small.

MUSICAL GESTURES

The predominant feeling that the *Presto* evokes is of headlong, forward motion. In addition to the fast tempo and 2/2 time of the notation, several features of the music contributed to this effect and these were given additional emphasis by the pianist. For example, the downward spike in bar 12, mentioned above (see Figures 2 & 3), corresponds to the downward octave jump that introduces the main (*A*) theme (see Figure 1); the upward spike that follows it corresponds to an ascending octave scale of eighth-notes in the right hand. The downward jump and the ascending scale in the score seem to reinforce the feeling of motion by giving the impression of a downward leap followed by a rush back up again. These musical images were accentuated by the pianist, who delayed the second note of the jump, which seemed to prolong the sensation of leaping down, and compressed the upward scale, which seemed to increase the feeling of rushing upward.

The upward and downward spikes in the temporal profile corresponding to the downward jump at the beginning of the *A* theme and the upward scale are indicated in Figure 4. Note that the size of the spikes varies, reflecting the fact that the same gesture was given varying amounts of emphasis at different points in the same performance. The pianist reported that she made a point of playing each repetition of the main theme somewhat differently and the variation in the gestures in Figures 2-4 confirms this.[5]

[5]The musical gesture Jump: *A* was identified in bars 1, 13, 65, 93, 113, and 199, Other Jumps in bars 23, 69, 75, 103, and 149, and Scales in bars 2, 12, 14, 64, 66, 92, 94, 114, 124, 140, 152, 186, 188, 198, 200.

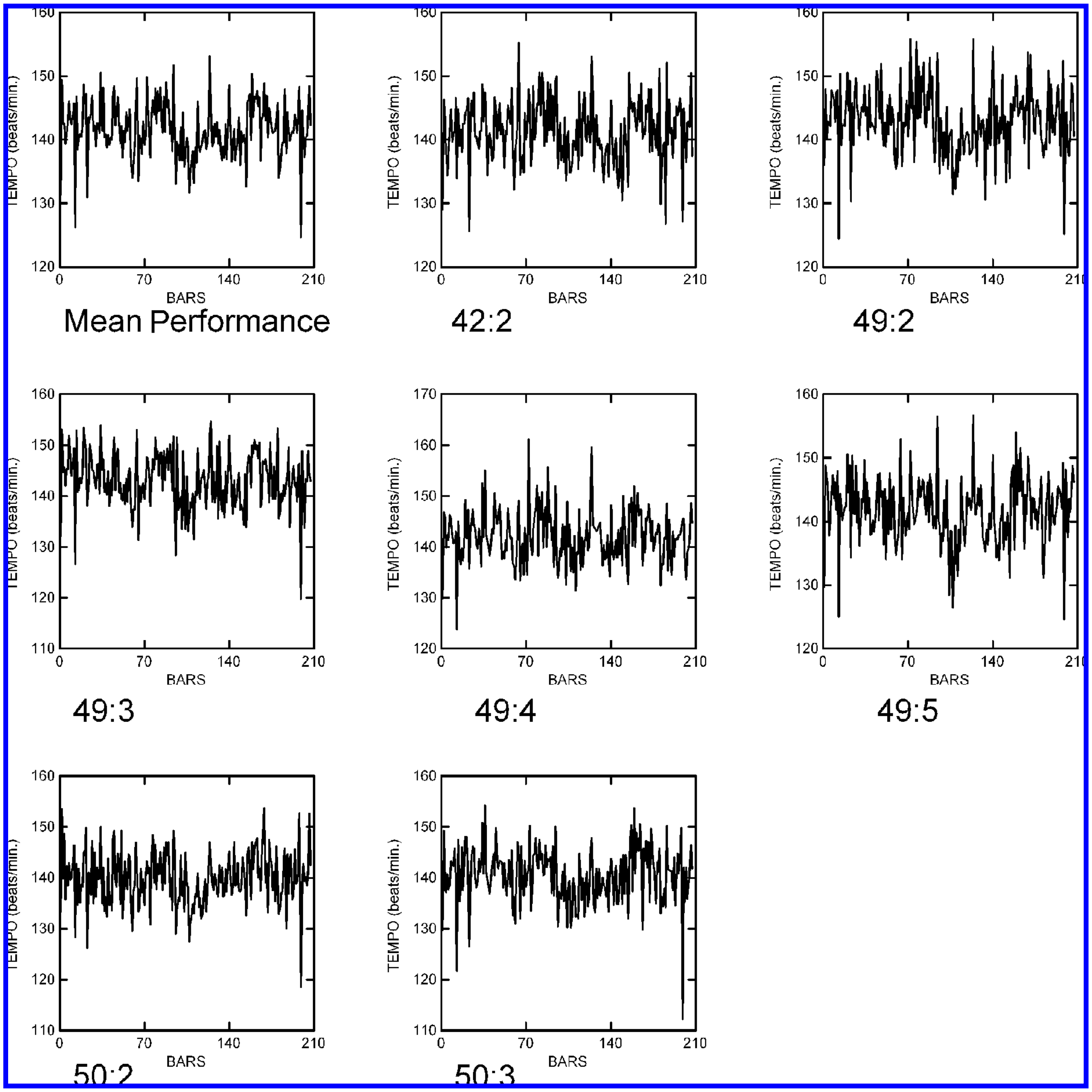

FIGURE 2. Tempo profiles of the mean performance and seven individual performances (last three bars omitted).

The accentuation of the downward octave jump at the start of the *A* theme was responsible for the negative (slowing) effect of jump: *A* on tempo in Table 4. The effect was significant for the mean performance and for all seven of the individual performances. Other large downward jumps at other points in the piece had a similar effect that was significant in the mean performance and in two of the individual performances. The accentuation of the upward rush of the ascending scale that often followed the jumps was the most consistent musical gesture. An increase in tempo added to the momentum of the upward rush, and the rapid succession of eighth notes produced an atypical consistency in dynamic level across the entire bar, resulting in lower dynamic variability. These effects of ascending scales on the mean performance were significant for all three measures and the effects were also significant for all seven of the individual performances for tempo and dynamic variability and for six of the seven performances for dynamic level.

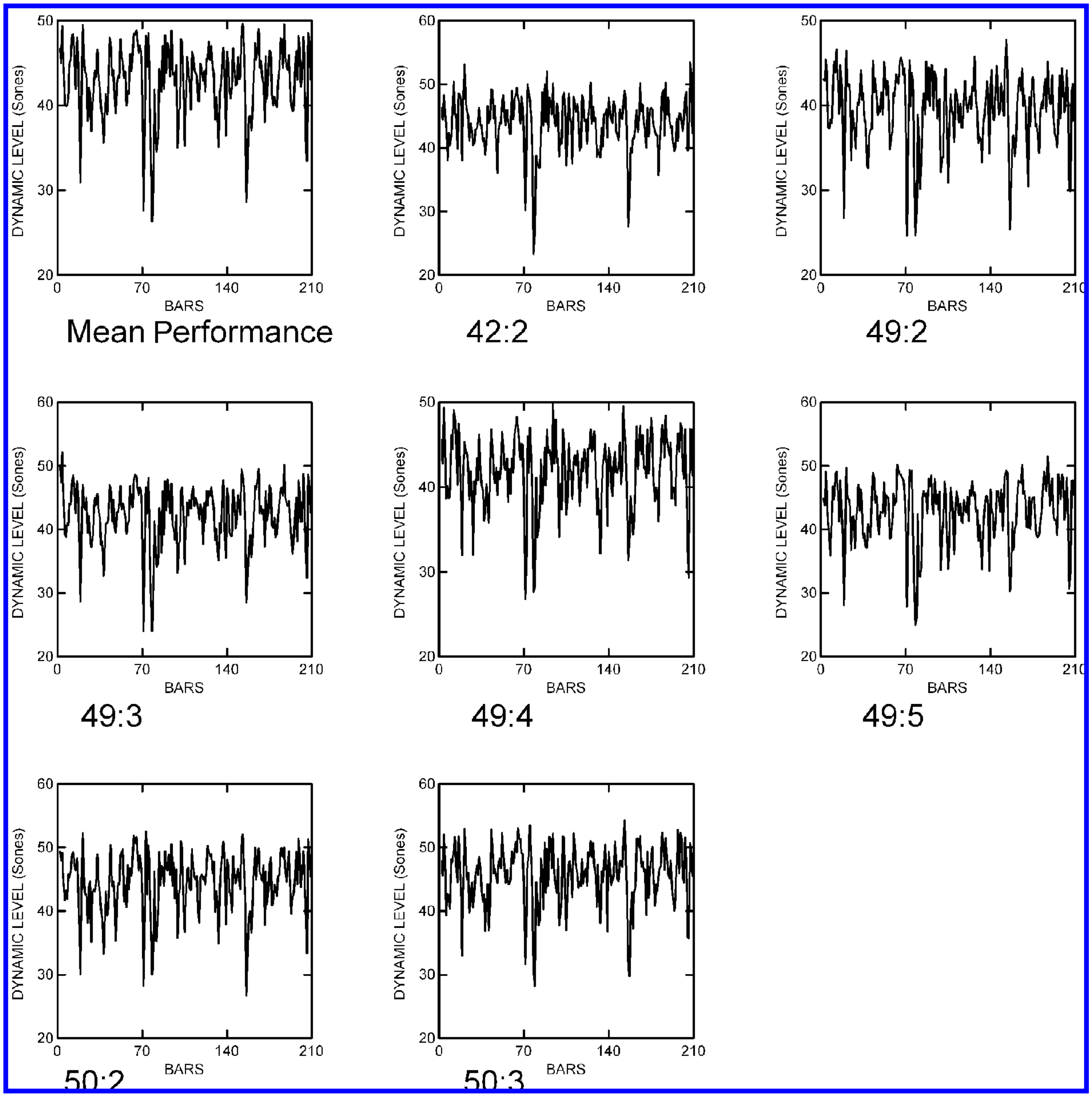

FIGURE 3. Dynamic level profiles of the mean performance and seven individual performances (first and last bars omitted).

Another musical feature of the *Presto* that contributes to the feeling of headlong motion is the insistent repetition of the main theme, which seems to give the impression of returning again and again to the same place as if on a wild carousel ride. The repetition is characteristic of the Italian rondo form of the *Presto*. A main theme (*A*) returns nine times separated by *B, C*, and *D* themes. This formal structure is reflected in the large-scale wave structure of the profiles (see Figures 2 & 3) and was reflected in four separate gestures. First, two sections containing the light, dancing *C* theme were faster and quieter (M = 145 beats/min, 40 sones) while the complex fugue in the *D* theme that marks the center of the piece was slower and louder (M = 136 beats/sec, 42 sones). In Table 4 these differences were responsible for the effects of sections on tempo and dynamic level that indicate that the differences between the themes were significant in the mean performance for both tempo and dynamic level and that these effects were also significant in six individual performances for

TABLE 4. Regression coefficients, R^2, and autocorrelation for effects of musical gestures and potential sources of stability on tempo, mean dynamic level, and dynamic variability for the mean performance.

		Dynamic level	
Predictors	Tempo	Mean	Variability
Musical gestures			
Jump: *A*	−8.718*** (7)	2.963 (2)	−0.831 (0)
Other jumps	−3.472* (2)	1.660 (0)	−1.589*** (0)
Scales	5.007*** (7)	2.865** (6)	0.045 (7)
Sections	−1.891*** (7)	0.720* (3)	−0.259 (1)
Section ends	−1.363* (1)	−0.622 (0)	0.674** (4)
Position in section	0.078 (2)	0.125 (0)	−0.129 (0)
Modulation	−1.062 (1)	0.446 (0)	−0.064 (0)
Phrasing	0.085 (1)	−0.201 (0)	0.166** (6)
Expressive intensity	0.137 (0)	1.462*** (7)	−0.064 (0)
Sources of stability			
Basic performance cues	−1.283*** (6)	0.389 (0)	−0.329* (4)
Technical difficulties	−0.959 (1)	1.201* (3)	−0.362 (1)
R^2	.459***	.395***	.232***
Auto-correlation	.075	.292*	−.058

Note: Number of significant effects in the 7 individual performances is included in parentheses. *p < .05, **p < .01, ***p < .001

tempo and in three individual performances for dynamic level.

A second gesture involving the formal structure is represented by the effects of ends of sections. The boundaries between sections were marked by a decrease in tempo in the last bar of the section and an increase in dynamic variability resulting from a momentary drop in dynamic level marking the transition to a new section. These effects were significant in the mean performance for tempo and dynamic variability, in one of the individual performances for tempo, and in four individual performances for dynamic variability. This kind of use of tempo and dynamics to delineate musical structure is a common feature of musical interpretation (Clarke, 1988, Gabrielsson & Lindström, 2001; Palmer, 1989, 1997; Sundberg, 1988). Two additional gestures involving musical structure affected individual performances without having significant effects on the mean performance. There was an effect of serial position in a section in two of the individual performances due to a gradual increase in tempo within sections, and of modulation on one performance due to slowing at transitions to a new key.

Phrasing increased dynamic variability in the mean performance and in six of the seven individual performances. The effect of phrasing was probably due to the use of short pauses to separate and shape phrases (De Poli, Roda, & Vidolin, 1998; Repp, 1998). The increase in dynamic variability could have been due to brief drops in dynamic level marking the start of a new phrase, similar to those between sections but affecting dynamic level rather than tempo. The increase in dynamic variability at the beginning of phrases was significant for the mean and

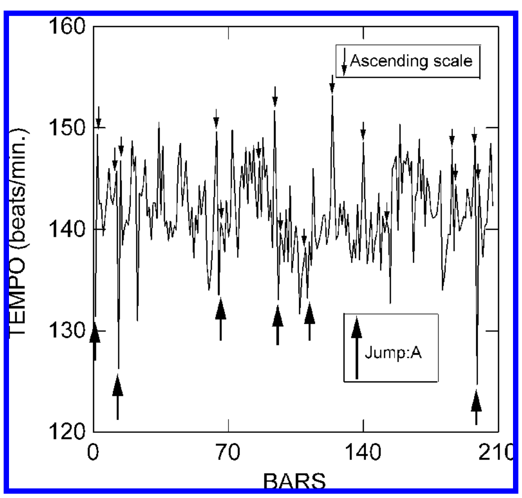

FIGURE 4. Tempo profile of the mean performance showing the location of two musical gestures introducing the *A* theme: The ascending 8th note scale and the octave jump down (Jump: *A*).

for six of the individual performances. Finally, higher levels of expressive intensity were marked by higher dynamic levels for the mean performance as well as for all seven of the individual performances. Similar use of increased dynamic level to express more active emotions has been observed in other performers (Juslin, 2000, 2003).

SOURCES OF STABILITY

Basic performance cues appeared to function as sources of stability, while technical difficulties did not. Table 4 shows that dynamic level was less variable in the mean performance for bars containing more basic performance cues. The effect was robust, occurring not only in the mean performance but also in four of the seven individual performances. The decreased variability was consistent with our suggestion that basic cues served as sources of stability. On this account, reducing the range of dynamic levels helped to achieve precision and reliability in the execution of actions that were essential for the performance to continue as planned. We suggest that the pianist's attention to the precision of movement diminished her ability to bring out nuances of phrasing and articulation, temporarily reducing variation in dynamic level. This does not mean that the decrease in interpretive nuances was musically inappropriate. Rather, the pianist's interpretive choices would have been shaped by multiple considerations including the demands on her attention.

There was a similar trend towards less dynamic variability in bars containing technical difficulties, but it was not significant for the mean performance. The weakness of this effect for technical difficulties suggests that the pianist had successfully mastered technical problems during practice so that they no longer affected her performance.

In addition to their effects on dynamic variability, tempo was slower at basic performance cues and dynamic level was higher at technical difficulties than at other points in the performance. The effects were very small, just over 1 beat/minute (4 ms) for basic cues and less than 1 sone for technical difficulties, and were not detectable by ear, at least by the authors. Taking a little more time at basic cues may have contributed to the overall character of the performances, which the pianist characterized as "cautious," attributing this to her desire for a note-perfect performance. The effect of technical difficulties on dynamic level may have been a result of the tendency commonly experienced by musicians to play difficult passages louder and may be due to a need to increase amplitude or force to overcome inertia in the motor system for difficult coordinations (Stims & Michaels, 1999).

ALTERNATIVE EXPLANATIONS

Could it be that the effects of musical gestures and sources of stability were due to musical properties other than those we have identified? To examine this possibility we re-ran the analyses, adding predictors for every property of the score that had been quantified in earlier studies that was not already included in the analyses reported above. The important effects remained unchanged. The effects of the musical gestures were still there, and dynamic variability was still lower and tempo slower at basic performance cues. These effects for tempo did not change when the dynamic mean profile was added as yet another additional predictor in the analysis of tempo. It is unlikely, therefore, that the differences between musical gestures and sources of stability were due to unidentified musical properties.

Basic performance cues were the only predictor related to technique to affect dynamic variability. The absence of effects for technical difficulties and fingering suggests that the effect for basic performance cues was not due to technique per se. The pianist was not just playing with less articulation in difficult passages. This strengthens our suggestion that that the effect was the result of the presence of basic performance cues at these spots, reflecting the pianist's decision to monitor these movements to ensure that they occurred as planned.

Another possible explanation for the effect of basic performance cues on dynamic variability was that the musical features that led the pianist to set up basic performance cues also called musically for less dynamic variability. If so, then the lower dynamic variability at basic cues may have been due, not to the extra attention paid to technique during performance as we have suggested, but to the hypothesized, unidentified, musical property. This explanation is made less plausible by the fact that the effect of basic performance cues on dynamic variability was still present in the additional analyses that included every aspect of musical structure, interpretation, performance, and technique that the pianist considered relevant. If any of these properties had been responsible for the effects we have reported, they would have disappeared when the additional predictors were added. They did not. While it is possible that some additional, unidentified, musical property was responsible for the effect, the presence of basic performance cues provides a plausible explanation.

Differences Between Performances

We identified differences between performances by comparing each individual performance with the mean performance. The procedure is conservative with

TABLE 5. Regression coefficients, R^2, and autocorrelation for the effects of musical gestures and potential sources of stability on the tempo of individual performances with effects present in the mean performance removed.

	Performance						
	1	2	3	4	5	6	7
Predictors	42:2	49:2	49:3	49:4	49:5	50:2	50:3
Mean performance	0.894***	1.031***	1.011***	0.975***	1.018***	1.082***	0.989***
Musical gestures							
Jump: *A*	1.114	0.083	−1.092	**2.869***	−0.072	0.322	**−3.224***
Other jumps	−2.352	−0.016	1.348	1.411	0.659	−1.728	0.678
Scales	0.793	0.068	−0.193	0.067	0.201	−1.455	0.519
Sections	−0.234	0.204	−0.365	−0.285	0.145	**0.767***	−0.234
Section ends	0.144	−0.032	0.134	−0.929	−0.053	1.011	−0.275
Position in section	**0.203****	−0.125	0.139	−0.053	**0.142***	**−0.185***	−0.120
Modulation	1.638	−0.563	−1.360	**3.740*****	**−1.569***	−0.551	−1.335
Phrasing	0.162	0.027	**0.388****	**−0.302***	−0.142	−0.046	−0.087
Expressive Intensity	0.054	0.015	0.037	−0.299	0.196	0.127	−0.130
Sources of stability							
Basic performance cues	0.005	0.097	−0.156	−0.148	−0.520	**0.790***	−0.067
Technical difficulties	−0.683	−0.086	−0.205	0.535	−0.048	0.033	0.455
R^2	.673***	.733***	.702***	.670***	.805***	.665***	.720***
Autocorrelation	−.084	−.119	−.228	−.116	−.149	−.112	−.185

*$p < .05$, **$p < .01$, ***$p < .001$

respect to finding differences between performances, because each performance was compared with itself as well as with other performances, but risks inflating the Type I error rate because each performance was analyzed separately. To protect against the latter possibility, preliminary omnibus *F* tests were performed for each measure (tempo, mean dynamic level, and dynamic variability). The seven performances were included as a repeated measures factor in mixed ANOVAs with musical gestures and sources of stability as independent variables, performances as the repeated measures factor, and bars as the replication (random) factor. Differences between performances in the effects of musical gestures and sources of stability were indicated by their interactions with performance. For tempo there were interactions with performance for jump: *A*, $F(6,1170) = 22.39$, $p < .04$, modulation, $F(6,1170) = 36.08$, $p < .002$, phrasing, $F(6,1170) = 25.27$, $p < .02$, and serial position in a section, $F(6,1170) = 36.09$, $p < .002$. There were no significant interactions for mean dynamic level, $F(6,1176) \leq 1.73$, $p > .10$, or for dynamic variability, $F(6,1176) \leq 2.02$, $p \geq .06$; therefore, subsequent comparisons of individual performances with the mean performance were not performed for these measures.

For tempo, each individual performance was compared with the mean performance in regression analyses similar to those described above but with the addition of the mean performance as an additional predictor (see Table 5). The regression coefficients represent the effects of each predictor when the tempo common to all performances was held constant. The analyses are thus equivalent to regression analyses of the differences between the mean performance and each of the individual performances (Campbell & Kenny, 1999).[6] The effect of the mean performance was significant for all of the individual performances, reflecting the high intercorrelations between performances, and between each performance and the mean performance.

In addition, there were ten significant effects for musical gestures and one for basic performance cues, indicating differences between the individual performances and the mean performance in the effects of these predictors. Nine of the effects for musical gestures involved the four gestures that interacted with performance in the ANOVA reported above and were, therefore,

[6]We included the mean performance as a predictor rather than computing differences so that in the subsequent stepwise analysis reported below we could assess the contribution of the mean performance along with other sources of variance to the individual performances.

reliable when the alpha level was controlled for the number of tests performed. The differences in Table 5 can only be viewed as reliable if they were also significant in the ANOVA that provided an omnibus test in which overall alpha level was controlled because the alpha level for the analyses in Table 5 was not controlled for the separate evaluation of each performance. Since the effects for sections and for basic performance cues were not significant in the ANOVA, their effects in the regression analyses summarized in Table 5 cannot be viewed as reliable.

There were nine reliable differences for four of the nine gestures and these occurred in six of the seven performances. Their prevalence suggests that, as we hypothesized, musically significant differences are a normal feature of repeated performances. The differences were not due to atypical performance of single bars; each gesture involved multiple bars and the differences, therefore, occurred consistently throughout the performance. Nor were the differences the result of random variability. Individual performances differed systematically from the mean performance profile in ways that were large enough to stand out from the background variability, thus the differences are unlikely to have been a product of noise alone. The statistical significance of the differences suggests that they were a result of more systematic processes. The effects are consistent with our suggestion that musically meaningful differences between repeated performances are an unavoidable byproduct of the psychological processes involved in playing musically. Musical gestures were more flexible than other aspects of performance, making it possible for repeated performances to differ from one another.

There were reliable differences between performances for four gestures. The slower tempo marking the octave jump at the beginning of the *A* theme was more pronounced in performance 50:3 and less pronounced in 49:4. Differences in three gestures involved the use of tempo to delineate phrases, sections, and key modulations. First, we have already suggested that boundaries between phrases were marked by brief pauses resulting in increased dynamic variability at the starts of phrases. The two differences for phrasing suggest that this gesture affected tempo differently in different performances. In performance 49:3, bars containing the beginnings of phrases were played more quickly than in the mean performance, while in 49:4 they were played more slowly. Second, the effect of serial position in a section for performances 42:2 and 49:5 indicates that, compared to the mean performance, sections were delineated more by a steadily increasing tempo profile in these performances. The negative effect in 50:2, in contrast, indicates that in this performance the tempo profiles showed less increase than in the mean performance. Third, the two effects for modulation indicate that modulation to a new key was marked by a greater decrease in tempo than in the mean performance for performance 49:5 and by a smaller decrease for performance 49:4.

Two remaining significant effects in Table 5 were not significant in the omnibus ANOVA and cannot be viewed as reliable. Both occurred in performance 50:2 and involved the accentuation of effects that were present in the mean performance. The difference in tempo between the four themes that determined the major sections of the piece and the slowing at basic performance cues were both less pronounced than in the mean performance.

ALTERNATIVE EXPLANATIONS

Is it possible that the difference between performances occurred because some passages were inherently more variable than others? Since larger values of any measure are inherently more variable, then louder or slower passages would be more likely to differ between performances than other passages. On this account, the artist's skill lies in locating musical gestures and basic performance cues at points where the inherent variability of the music is higher or lower. To test these possibilities, the analyses were rerun with the number of notes per bar and with the maximum dynamic level in a bar included as additional predictors to control for length of note and dynamic level respectively.[7] The number of significant differences was unchanged, suggesting that inherent variability was not responsible for the differences.

Inherent variability is also implausible as an explanation for other reasons. First, there were differences between performances for the gesture involving the upward eighth-note scale which was faster (and thus inherently *less* variable) than other passages. Second, the inherent variability account predicts that variation in a gesture at one point in the performance would be unrelated to variation in the same gesture at another point. What we found, however, was that some gestures were systematically larger or smaller across an entire performance. These were not random changes in isolated passages but systematic changes in the way gestures were implemented throughout the piece.

[7]Maximum dynamic level was determined by taking the maximum value of the 10 ms slices of the sound signal for each bar.

Sources of Variation in the Performances

To find out how similar the performances were and how much they differed at sources of stability and at musical gestures, the regression analyses described in the previous section were repeated stepwise, entering the mean performance first, then the predictors representing potential sources of stability, and then the musical gestures (see Table 6). The individual performances were all very similar to the mean performance. In the first step, the mean performance accounted for 62-78% of the variation in tempo and 77-91% of the variation in dynamic level, with means of 68% and 87% respectively (see Table 6, column 1). Most of the remaining variability was due to noise. The greater similarity for dynamic level was probably due to its lower noise level (see below).

The performances were more similar at sources of stability than at musical gestures. Deviations from the mean performance at sources of stability were minimal (0.3% and 0.1% for tempo and dynamic level respectively). Deviations in the musical gestures, while also small in absolute terms, were much larger in relative terms (2.4% and 0.6% for tempo and dynamic level respectively). The difference is consistent with our hypothesis that variation is restricted at sources of stability and allowed to persist at musical gestures. The values for musical gestures also provide a quantitative lower bound for differences between performances. Even when conditions favored repetition of exactly the same performance, there was a 2% difference in tempo between performances in the musical gestures that we identified.

The remaining variation can be attributed to noise, which accounted for 29% and 11.8% of the variation respectively for tempo and dynamic level. Half of the noise for tempo can be attributed to error of measurement due to the difficulty of judging precisely where each bar began. A more direct measure of noise level for tempo is provided by the mean square error for the final step of the regression analyses: 8.6 beats/minute and 2.6 sones, averaged across performances. At a tempo of 141 beats/minute, this estimate for tempo is approximately 25 ms, which is double the mean error of measurement reported above in Table 1. It appears that for tempo approximately half of the noise was due to error of measurement (12 ms) with the other half (approximately 15% of the overall variance) attributable to real variation in the performances due both to random processes and to systematic but unidentified sources of variation. Dynamic level was less subject to measurement error. Errors in judgment about where bars started and stopped affected only a small proportion of the 10 ms slices that made up the dynamic measurements for each bar. The lower level of measurement error was reflected in the smaller proportion of variance attributed to noise in the dynamic level analyses (11.8%).

TABLE 6. Change in R^2 in stepwise regressions due to the mean performance, and to predictors representing potential sources of stability and musical gestures (interpretation) for tempo and mean dynamic level in individual performances.

		Performance						
		1	2	3	4	5	6	7
Source of variation	Mean % variance accounted for across performances	42:2	49:2	49:3	49:4	49:5	50:2	50:3
Step 1: Mean performance								
Tempo	68.4	.644	.726	.679	.620	.780	.635	.702
Dynamic level: Mean	87.5	.770	.888	.908	.881	.895	.871	.910
Step 2: Sources of stability								
Tempo	0.30	.004	.000	.001	.002	.004	.008	.002
Dynamic level: Mean	0.10	.004	.001	.001	.000	.001	.000	.001
Step 3: Musical gestures								
Tempo	2.40	.025	.007	.022	.053	.021	.022	.016
Dynamic level: Mean	0.60	.015	.005	.005	.007	.004	.003	.002
Unexplained variation (noise)								
Tempo	29.0	.327	.267	.298	.325	.195	.335	.280
Dynamic level: Mean	11.8	.211	.106	.086	.112	.100	.126	.087

Note: Mean changes in R^2 across performances are expressed as the percentage of variance accounted for.

Conclusion

Emil Gilels was right: Every performance is different. There were nine systematic differences in the musical gestures of the seven performances. The differences were few in number, occurred only for tempo, and were minor in comparison to the similarities and to the variation that was not explained; the musical gestures accounted for less than 3% of the overall difference in tempo and for less than 1% for dynamic level. However, the differences in tempo were pervasive, occurring in 6 of the 7 performances and for 4 of the 9 musical gestures; and their statistical reliability indicates that they were consistent. The differences were not limited to a single passage but occurred throughout the *Presto* as the pianist systematically exaggerated or minimized particular gestures throughout entire performances. The differences were a product of systematic processes, not just random variation.

The presence of reliable differences in these performances is remarkable, given that we looked for differences under conditions where they would be least expected. Preparation of the *Presto* was complete after almost 33 hours of practice over almost 10 months, and the pianist was striving to reproduce the same performance each time. All of the usual motivations for adjusting a performance to the characteristics of instrument, hall, and audience were absent; the performances were all in the artist's practice studio and for the same audience: the video camera. Six of the performances were recorded on the same day when the pianist was striving to videotape a "perfect" performance to play during talks about the research. In addition, the fast tempo and *perpetuo mobile* style of the *Presto* call for a steady, rhythmic beat with little dynamic variation, and performance conventions of the Baroque era require that changes in tempo or dynamics take place in discrete steps rather than varying continuously, as in the later Romantic style. Everything about these performances conspired to make them as similar as possible. The presence of differences under these conditions suggests that such differences are a normal product of the psychological mechanisms responsible for performance.

It is not important for our purposes whether the differences were audible to listeners or not, or whether the performer was aware of them. The purpose of musical gestures is to affect the listener; however, our purpose was not musical but scientific: to understand the motor and cognitive system that produces the gestures. We suspect that although most listeners might be able to detect differences when hearing selected passages, they would not remember them well enough to distinguish one performance from another. The pianist herself considered the performances to be very similar. This similarity is an advantage for our purposes. The kinds of differences we found are likely to be present in any other set of performances. We believe they would be more pronounced in live performances and with slower music that required more expressive variation (Repp, 1995).

The differences between performances were not random, but they also do not appear to have been deliberate. The variation *within* a performance in the gestures at the beginning of the *A* theme was deliberate, at least according to the pianist's report. The variation of gestures *across* performances, on the other hand, appears to have been less so and was probably a product of both random and purposeful influences. For example, it is possible that the emphasis given to a gesture the first time it occurred was determined randomly and that this then influenced the way that the pianist played succeeding instances of the same gesture.[8] For example, if the prolongation of the downward octave jump in the first bar was more pronounced than usual, then all of the following jumps at repetitions of the *A* theme would also be more pronounced. The calibration necessary to produce this kind of consistency within a performance can be described in terms of control parameters that establish a state-space within which each gesture is free to vary (Latash et al., 2002) or by selective, goal-directed editing to enhance a particular gesture throughout one performance and reduce it in another (Todorov & Jordan, 2002).

It is probably not fruitful to ask whether this kind of parameter setting or editing is deliberate or not (Dennett, 1991). On the one hand, the pianist's opinion that the performances were very similar and the fact that complex decisions are generally better when made unconsciously (Dijksterhuis & Nordgren, 2006) suggest that the pianist may not have been aware of the differences between performances. On the other hand, the pianist's main concern in sessions 49 and 50 was with accuracy and she may have considered the expressive differences between the performances too minor to be worth mentioning.

Our method was correlational and so it is possible that the effects that we found were not due to the musical gestures or points of stability we have identified, but to other properties of the music. Although we cannot entirely rule out such explanations, the additional analyses we performed tested the most obvious of these alternative explanations and found no support for them.

[8]We thank Bruno Repp and an anonymous reviewer for suggesting this explanation.

Passages containing musical gestures did not seem to be inherently more variable than other passages and the lower dynamic variability at basic performance cues did not seem to be due simply to technique or to any of the other properties identified by the pianist.

The differences may be an example of the kind of trade-off between stability and flexibility that is characteristic of the motor system at every level and is exploited by expert performers in other fields (Bernstein, 1967; Latash et al. 2002; Lee at al., 1982; Todorov & Jordan, 2002). It remains for future work to show that the flexibility and stability we have identified involved trade-offs. Here we have demonstrated that both were present in the same performances. When the pianist needed to attend to precision of execution, she set up basic performance cues by training herself to monitor critical movements (Chaffin et al., 2002, pp. 179-190). In contrast, when the pianist made musical gestures, she acted more flexibly, with the consequence that the gestures differed from one performance to the next. The differences were not due to a simple trade-off in which a general increase in variability was traded for reduced variability at basic performance cues. Rather, increases and decreases in variability were targeted specifically and increases were consistently calibrated throughout a performance. This is where the pianist was different from the marksmen and long-jumpers studied by Scholz et al. (2002) and Lee et al. (1982). The sportsmen in these studies managed variability to achieve precision of technique: variation in a long-jumper's stride approaching the plate serves no goal other than to optimize the take-off. Our pianist, in contrast, managed the variability in her performances both to achieve technical precision and to serve her aesthetic goals.

What aesthetic goal would be served by performances that were more flexible at musical gestures? Flexibility at these points may give a performance a freshness and spontaneity of musical expression that would otherwise be lacking in a piece that has been thoroughly prepared. Performers use a variety of strategies to deal with the dilemma of spontaneity in an over learned task (Chaffin et al., 2002, pp. 60-63). Pianist Stephen Bishop-Kovacevich finds that, "*It's a good idea not to touch any of the pieces you are due to play [on the day of the concert]*" (Dubal, 1997, p. 56), a strategy also used by the pianist in our study (Chaffin et al., 2002, p. 6). We suggest that flexibility of musical gestures is another strategy. Our results suggest that the pianist managed the variability in her performance by allowing more variability at musical gestures, and by exercising more control at basic performance cues where precision was needed for technical reasons.

Our results provide the first direct evidence that basic performance cues are used to control critical details of technique during polished performance. Previous work has shown that the pianist in the present study practiced these cues during much of her 33 hours of practice of the *Presto*, that she hesitated at these cues when she began to play from memory, and that her recall of the score was poorer at these points (Chaffin & Imreh, 2002; Chaffin et al., 2002). Other experienced soloists and conductors also systematically rehearse their use of basic performance cues (Chaffin, Lisboa, Logan, & Begosh, 2004; Chaffin & Logan, 2006; Ginsborg, Chaffin, & Nicholson, 2006; Noice, Chaffin, Noice, & Jeffrey, in press). Here we have shown, in addition, that slowing at basic cues continued when the music was thoroughly learned, and that dynamic variability decreased at the same points. We suggest that these effects show that basic performance cues are a source of stability. The pianist took more time at basic cues because she was making sure that critical details of technique were executed as planned and dynamic variability decreased because she had less attention to devote to nuances of interpretation.

Our study involved a single musician; studies of other musicians are needed to establish the generality of our conclusions. Case studies are the appropriate method because the 10-20 years of training required for a career as a soloist in the Western art music tradition increases the normal range of individual differences so that aggregating observations across individuals runs the risk of obscuring the phenomena of interest (Ericsson & Oliver, 1988). Generalization from case studies of exceptionally skilled individuals must be based on support for general psychological principles. The pianist in the present study practiced basic performance cues in the same way as expert musicians in other case studies (Ginsborg et al., 2006; Lisboa et al., 2006; Noice et al., 2006) and managed the trade-off between precision and variability in the same way as experts in other domains (Latash et al., 2002; Lee at al., 1982; Todorov & Jordan, 2002). These convergent findings suggest that other experienced musicians may use similar strategies to keep their performances accurate while maintaining freshness and spontaneity.

Author Note

We thank Gabriela Imreh for providing the impetus for this research and the performances analyzed here, Reinhard Kopiez for assistance in measuring dynamic levels, David Kenny and Len Katz for statistical advice, and Mary Crawford, Claire Michaels, Bruno Repp, and Michael Turvey for helpful suggestions on earlier versions of the paper. Preliminary analyses of these data were reported by Chaffin, Lemieux, & Chen (2004).

Correspondence concerning this article should be addressed to Roger Chaffin, Department of Psychology U-1020, University of Connecticut, Storrs, CT 06269-1020 USA; E-MAIL: *Roger.Chaffin@Uconn.edu*

References

BERNSTEIN, N. A. (1967). *Coordination of movement*. Oxford: Pergamon Press.

CAMPBELL, D. T., & KENNY, D. A. (1999). *A primer on regression artifacts*. New York: Guilford Press.

CHAFFIN, R., & IMREH, G. (1997). "Pulling teeth and torture": Musical memory and problem solving. *Thinking and Reasoning: Special Issue on Expert Thinking, 3*, 315-336.

CHAFFIN, R., & IMREH, G. (2001). A comparison of practice and self-report as sources of information about the goals of expert practice. *Psychology of Music, 29,* 39-69.

CHAFFIN, R., & IMREH, G. (2002). Practicing perfection: Piano performance as expert memory. *Psychological Science, 13*, 342-349.

CHAFFIN, R., IMREH, G., & CRAWFORD, M. (2002). *Practicing perfection: Memory and piano performance*. Mahwah NJ: Erlbaum.

CHAFFIN, R., IMREH, G., LEMIEUX, A. F., & CHEN, C. (2003). "Seeing the big picture": Piano practice as expert problem solving. *Music Perception, 20*, 461-485.

CHAFFIN, R., LEMIEUX, A. F., & CHEN, C. (2004, August). *"It's different each time I play": Why highly polished performances vary*. Paper presented at the 8th Annual International Conference on Music Perception and Cognition, Evanston, IL.

CHAFFIN, R., LISBOA, T., LOGAN, T., & BEGOSH, K. (2006, May). *Expert memory in solo cello performance*. Poster presented at the annual meeting of the Association for Psychological Science, New York, NY.

CHAFFIN, R., & LOGAN, T. (2006). Practicing perfection: How concert soloists prepare for performance. *Advances in Cognitive Psychology, 2,* 113-130.

CLARKE, E. F. (1988). Generative principles in music performance. In J. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 1-26). New York: Oxford University Press.

CLARKE, E. F. (1995). Expression in performance: Generativity, perception and semiosis, In J. Rink (Ed.), *The practice of performance* (pp. 21-54), Cambridge: Cambridge University Press.

CLARKE, E. F. (2004). Empirical methods in the study of performance. In E. Clarke & N. Cooke (Eds.), *Empirical musicology: Aims, methods, prospects* (pp. 77-102). New York: Oxford University Press.

CLYNES, M., & WALKER, J. (1982). Neurobiologic functions of rhythm, time and pulse in music. In M. Clynes (Ed.), *Music, mind and brain: The neuropsychology of music* (pp. 171-216). New York: Plenum Press.

CLYNES, M., & WALKER, J. (1986). Music as time's measure. *Music Perception, 4*, 85-120.

DENNETT, D. (1991). *Consciousness explained*. Boston: Little Brown & Co.

DEPOLI, G., RODA, A., & VIDOLIN, A. (1998). Note-by-note analysis of the influence of expressive intentions and musical structure in violin performance. *Journal of New Music Research, 27*, 293-321.

DIJKSTERHUIS, A., & NORDGREN, L. F. (2006). A theory of unconscious thought. *Perspectives on Psychological Science, 1*, 95-109.

DUBAL, D. (1997). *Reflections from the keyboard: The world of the concert pianist* (2nd ed.). New York: Schirmer.

ERICSSON, K. A., & OLIVER, W. (1988). Methodology for laboratory research on thinking: Task selection, collection of observation and data analysis. In R. J. Sternberg & E. E. Smith (Eds.), *The psychology of human thought* (pp. 392-428). Cambridge: Cambridge University Press.

GABRIELSSON, A. (1999). The performance of music. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 579-602). San Diego: Academic Press.

GABRIELSSON, A., & LINDSTRÖM, E. (2001). The influence of musical structure on emotion. In P. N. Juslin & J.A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 223-248). New York: Oxford University Press.

GILDEN, D. L. (2001). Cognitive emissions of 1/f noise. *Psychological Review, 108*, 33-56.

GINSBORG, J., CHAFFIN, R., & NICHOLSON, G. (2006). Shared performance cues in singing and conducting: A content analysis of talk during practice. *Music Psychology, 34*, 167-194.

HINSON, M. (1987). *Guide to the pianist's repertoire* (2nd ed.). Bloomington: Indiana University Press.

IMREH, G. (PIANIST). (1996). *J. S. Bach* [CD]. New York: Connoisseur Society.

IMREH, G., & CHAFFIN, R. (1996/97). Understanding and developing musical memory: The view of a concert pianist. *American Music Teacher, 46*, 20-24, 67.

JUSLIN, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1797-1813.

Juslin, P. N. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129*, 770-814.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology, 51*, 336-352.

Latash, M. L., Scholz, J. P., & Schöner, G. (2002). Motor control strategies revealed in the structure of motor variability. *Exercise and Sports Science Reviews, 30*, 26-31.

Lee, D. N., Lishman, J. R., & Thomson, J. A. (1982). Regulation of gait in long jumping. *Journal of Experimental Psychology: Human Perception & Performance, 8*, 448-459.

Mach, E. (1991). *Great contemporary pianists speak for themselves*, Vol. 2. New York: Dover.

Newell, K. M., & Corcos, D. M. (1993). Issues in variability and motor control. In K. M. Newell & D. M. Corcos (Eds.), *Variability and motor control* (pp. 1-12). Champaign, IL: Human Kinetic.

Noice, H., Chaffin, R., Noice, A., & Jeffrey, J. (in press). Memorization by a jazz pianist: A case study. *Music Perception*.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 331-346.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48*, 115-138.

Repp, B. H. (1992). Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's Träumerei. *Journal of the Acoustical Society of America, 92*, 2546-2568.

Repp, B. H. (1995). Quantitative effects of global tempo on expressive timing in music performance: Some perceptual evidence. *Music Perception, 13*, 39-57.

Repp, B. H. (1998). A microcosm of musical expression: I. Quantitative analysis of pianist's timing in the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America, 104,* 1085-1100.

Ross, A. (2005, June 6). The record effect: How technology has transformed the sound of music. *The New Yorker,* 94-100.

Scholz, J. P., Schöner, G., & Latash, M. L. (2000). Identifying the control structure of multijoint coordination during pistol shooting. *Experimental Brain Research, 135,* 382-404.

Shaffer, L. H. (1984). Timing in solo and duet piano performances. *Quarterly Journal of Experimental Psychology, 36A,* 577-595.

Shaffer, L. H., Clarke, E. F., & Todd, N. P. (1985). Meter and rhythm in piano playing. *Cognition, 20*, 61-77.

Stims, J. F., & Michaels, C. F. (1999). Stimulus-response vs. stimulus-manipulandum compatibility in rhythmic coordination of oscillatory movements. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1793-1812.

Sundberg, J. (1988). Computer synthesis of music performance. In J.A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 52-69). New York: Oxford University Press.

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience, 5,* 1226-1235.

Zwicker, E., & Fastl, H. (1999). *Psychoacoustics. Facts and models* (2nd ed). Berlin: Springer.