A microcosm of musical expression: II. Quantitative analysis of pianists' dynamics in the initial measures of Chopin's Etude in E major

Bruno H. Repp[a)]
Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511-6695

(Received 4 May 1998; revised 9 November 1998; accepted 12 November 1998)

Patterns of expressive dynamics were measured in bars 1–5 of 115 commercially recorded performances of Chopin's Etude in E major, op. 10, No. 3. The grand average pattern (or dynamic profile) was representative of many performances and highly similar to the average dynamic profile of a group of advanced student performances, which suggests a widely shared central norm of expressive dynamics. The individual dynamic profiles were subjected to principal components analysis, which yielded five Varimax-rotated components, each representing a different, nonstandard dynamic profile associated with a small subset of performances. Most performances had dynamic patterns resembling a mixture of several components, and no clustering of performances into distinct groups was apparent. Some weak relationships of dynamic profiles with sociocultural variables were found, most notably a tendency of female pianists to exhibit a greater dynamic range in the melody. Within the melody, there were no significant relationships between expressive timing [Repp, J. Acoust. Soc. Am. **104**, 1085–1100 (1998)] and expressive dynamics. These two important dimensions seem to be controlled independently at this local level and thus offer the artist many degrees of freedom in giving a melody expressive shape. © *1999 Acoustical Society of America.* [S0001-4966(99)00803-6]

PACS numbers: 43.75.St, 43.75.Mn [WJS]

INTRODUCTION

In a previous article, Repp (1998a) reported a detailed quantitative analysis of pianists' expressive timing (the pattern of inter-onset interval durations between successive tones) in bars 1–5 of 115 commercially recorded performances of Chopin's Etude in E major, op. 10, No. 3. Although the grand average timing pattern was representative of many performances, principal components analysis of the data suggested at least four independent strategies of deviating from this central norm. Each individual pianist's timing pattern could be approximated by a weighted combination of these four strategies. A wide variety of combinations was represented in the sample, and no two individual patterns were exactly the same. In addition, there was a wide range of basic tempi and of degrees of tempo modulation. There were no strong relationships between any of these variables and sociocultural characteristics of the artists, although some weak trends were observed.

Timing is only one aspect of musical expression, though a very important one. It also happens to be the one that is easiest to measure in acoustically recorded performances. Another extremely important aspect of expression on the piano is dynamics—the relative intensities of successive and simultaneous tones. Although musical-instrument-digital-interface (MIDI) recordings give easy access to this information in the form of key-press velocities that are highly correlated with the sound levels of tones, expressive dynamics has received much less attention than expressive timing in performance research (but see Shaffer, 1981; Gabrielsson, 1987; Todd, 1992; Repp, 1996a). In particular, very little is known objectively about individual differences in expressive dynamics, even though informal listening suggests that such differences do exist in performances of the same music by great artists, preserved in acoustic recordings.

One methodological problem in measuring the expressive dynamics of acoustic recordings is that it is extremely difficult to estimate accurately the relative intensities of several simultaneous complex tones, even when their fundamental frequencies are known. Their harmonics are interleaved and often coincident, unpredictable phase relationships affect the amplitude at any given frequency, and the amplitude of the lowest harmonic (the fundamental frequency) is only roughly proportional to a tone's overall amplitude (Repp, 1993). At present, there seems to be no signal processing algorithm that can perform this task. Therefore, analyses of expressive dynamics in acoustically recorded performances (such as Gabrielsson, 1987, and the present study) are currently restricted to measuring the overall amplitudes of successive tone clusters—the "horizontal" dynamics. Information about "vertical" dynamics (the relative intensities of simultaneous tones, constituting the sonic texture) remains unavailable.

It is possible to ignore the fact that most successive acoustic events in music are composed of several individual tones and to regard total sound energy as the perceptually most relevant measure of horizontal dynamics (as does Todd, 1992, 1994). This may be sufficient for perception of rhythm. However, human listeners perceptually segregate music into individual voices if the compositional structure provides a basis for doing so. In homophonic music, as used in the present study, one voice (the melody, usually in the

[a)]Electronic mail: repp@haskins.yale.edu

highest pitch register) is more important than the others (the accompaniment) and is usually played with greater intensity. The listener's attention is drawn to the melody, and therefore the expressively most relevant measure of horizontal dynamics would seem to be the intensities of the melody tones (cf. Palmer, 1996a). From this perspective, measurements of overall intensity are but an approximation to the desired measure of horizontal dynamics. They seem to provide a sufficiently informative approximation, however, and will have to do until more sophisticated signal processing algorithms are developed.

This article presents the results of detailed analyses of the horizontal dynamic patterns measured in the 115 recordings of the Chopin Etude excerpt, with special attention to individual differences. The analyses were largely analogous to those of expressive timing presented in Repp (1998a). The main question was whether, and how many, fundamentally different patterns of dynamic shaping can be distinguished in this large sample of performances, and how these patterns are related to the musical structure. A secondary question was whether the identified patterns have any relationship to sociocultural differences among the artists.

Towards the end of the article, possible relationships between expressive timing and dynamics are examined. This is an extremely important issue about which very little is known at present. Todd (1992) has pointed out a tendency for dynamics to increase along with tempo and thus to be inversely related to expressive timing (measured in terms of inter-onset intervals). This is only a global tendency, however, that is often violated. Over longer musical passages, a significant correlation is usually obtained because the large-scale phrase structure constrains both timing and dynamics (Todd, 1992; Palmer, 1996a; Repp, 1996a). However, it is not clear whether a similar relationship holds at the most detailed level of expressive variation within a phrase, which is the subject of the present study (see also Palmer, 1996a). Also, given that substantial individual differences exist in both timing and (as the present data will show) dynamics, the question arises whether these individual differences covary. Repp (1996a) addressed this issue previously in performances of Schumann's "Träumerei" by ten advanced student pianists who did not show very large individual differences in either timing or dynamics. He did find a weak but significant relationship between the residual patterns (deviations from the average pattern) of timing and dynamics, such that some pianists tended to play more softly than average when they played slower than average at any given moment in the music. In the present sample of distinguished recording artists, much larger individual differences were expected, which provided an opportunity to reexamine the relationship of timing and dynamics.

I. METHOD

A. The music

A simplified score of the musical excerpt is shown on top of Fig. 1.[1] The music is divided into several horizontal strands or voices. The melody is in the soprano voice and includes both short (sixteenth) and longer notes. The accompanying alto voice has sixteenth notes throughout. Some additional filler notes in the right-hand part (which may be considered as forming a "mezzo-soprano" voice) and the lower voices (tenor and bass) in the left-hand part also have an accompanying function. The melody is clearly the most important voice and is generally also played with greater intensity than the other voices. When melody tones are sustained, the alto (bars 1–3) and mezzo-soprano/alto (bar 4) voices come to the fore and are more important than the lower voices. Following Repp (in press), the 38 alternating soprano and alto or mezzo-soprano notes forming the "top line" of the musical texture will be referred to as *primary notes* (or primary tones, when played). There is one primary note in every sixteenth-note metrical position of the music (for simplicity, the initial eighth-note upbeat will be treated here as if it were another sixteenth note), and it is defined to be always the highest note in its position. (Thus, in bar 4, positions 6–8, the mezzo-soprano notes are the primary notes.) Only four primary notes are not accompanied by lower note onsets: the initial upbeat and the alto notes in the seventh metrical position of bars 1, 2, and 3. These alto notes, however, are accompanied by sustained notes in both higher and lower voices. Since piano tones decay over time (see Martin, 1947; Repp, 1997c), these sustained tones have lost some of their energy when a primary alto note is played.

B. Measurement and analysis procedures

It was assumed that the perceptually most relevant horizontal dynamic pattern is that of the primary tones. It reflects the dynamic shaping of the melody as well as the dynamic contrast between melody and accompaniment. However, since the relative intensities of the primary tones could not be measured directly, the question arose to what extent the pattern of overall intensities resembled that of the primary tones alone. Clearly, additional tones accompanying a primary tone will raise the overall intensity, and the extent of that increase is likely to depend on the number and relative intensity of these additional tones. The acoustic complexity of piano tones and of their temporal relationships in expressive performance makes it difficult, however, to predict these effects mathematically. Therefore, some preliminary measurements and analyses were performed on synthesized performances, to get some indication of the extent to which the overall dynamic profile parallels the dynamic profile of the primary tones. These results are described in Appendix A.

A total of 117 recordings of the Chopin excerpt were procured. (For a detailed description, see Repp, 1998a.) Most of them (102) were obtained on a digital audio tape onto which they had been copied from compact discs and long-playing records. (Two of these copies turned out to be duplicates of others, but they were included in all analyses.) Although there was considerable variation in the recorded sound level, no distortion was noticed when listening to the tape. The remaining 15 recordings were available from a previous study (Repp, 1997a), saved as digitized sound files. Using SOUNDEDIT16 software, all recordings were input as analog signals to a Macintosh Quadra 660AV computer, sampled at a rate of 20.055 kHz, and stored as separate files in 16-bit format. (Some of the 15 recordings digitized earlier were in 8-bit format.) Each waveform was subsequently scaled multiplicatively to maximum peak amplitude using the "normalize" function of SOUNDEDIT16. Measurements of tone (chord) onset times had already been performed (Repp, 1998a) and served to guide the amplitude analysis.

Using SIGNALYZE software, the root-mean-square amplitude envelope of each digitized waveform was computed with a rectangular integration window of 30 ms, and subsequently an automatic peak-picking routine was employed to determine the peak amplitude following each note (chord) onset. The amplitudes were then converted into peak sound levels (PSLs) in decibels (dB). These measurements obviously did not achieve the precision and validity of the previous timing measurements (Repp, 1998a). Interactions among simultaneous tones and multiple sources of distortion (surface noise, recording techniques, room acoustics) made the data rather noisy. Nevertheless, they were believed to be sufficiently informative about the pianists' dynamic strategies to justify the following analyses.

Each performance yielded a series of 38 PSL values (cf. the score in Fig. 1) that constituted its *dynamic profile*. These data were subjected to various correlational analyses, including principal components analysis with Varimax rotation. A brief introduction to this technique may be found in Repp (1998a). Twenty-three of the 38 PSLs derived from positions in which there were note onsets in the soprano voice. Separate analyses were conducted on these *melodic dynamic profiles*, in order to separate melodic dynamic variation from the usually large dynamic difference between melody and accompaniment.

II. RESULTS AND DISCUSSION

A. Three aspects of expressive dynamics

Like expressive timing, expressive dynamics has three largely independent aspects. The first of these is the *basic dynamic level*, which can be measured by computing the average PSL of a whole passage and which corresponds to dynamic markings in the score (such as *piano* or *forte*) that apply across a number of bars. In the present analyses, however, the average PSL was meaningless because it depended on several uncontrolled factors that varied between performances, including the original recording level, the recording level in the transfer from the original medium to digital tape, the input level to the computer, and the subsequent normalization to maximum amplitude. Although human listeners may be able to recover information about the dynamic level of the original performance from the spectral content of piano tones, no spectral or perceptual analyses were attempted in the present study. All absolute dB values in the figures displayed in this article are meaningless.

The second aspect of expressive dynamics is its *variability* or *range*, which is most conveniently measured by the standard deviation of the PSLs or a multiple thereof. This measure is meaningful in the present context, even though it represents a conflation of the pianist's dynamic range in performance and the dynamic range of the recording. Old recordings in particular may exhibit a restricted dynamic range, even though the original performance may have had a wide range. High surface noise from old recordings may also reduce the measured dynamic variability by masking soft tones.

The third aspect is the pattern of PSL values that constitutes the *dynamic profile*. This was the aspect of primary interest in this study, and it was the only aspect that entered the correlational statistics because they entailed a conversion of the data into standard scores having a mean of zero and a standard deviation of one. Therefore, the findings on dynamic variability will be discussed first.

B. Dynamic variability and range

Although the difference between the largest and smallest PSL values in a performance would be the most direct measure of its dynamic range, the standard deviation (s.d.) provides a more robust index. The absolute dynamic range may be taken to be approximately four times the s.d. (i.e., the width of a 95% confidence interval around the mean PSL). The performance with the widest dynamic range was that by Cherkassky (s.d.=6.8 dB, a range of about 27 dB), followed by Kyriakou (6.5 dB), and Duchâble (5.9 dB). The performances with the narrowest dynamic range were those by W. Haas, Malcuzinsky, and Horowitz-1972 (all s.d.'s=2.4 dB, a range of about 10 dB).

In the present musical excerpt, the overall dynamic variability reflects in part the difference in average dynamic level between melody and accompaniment, that is, the difference between the 23 positions in the music in which there is a melody note onset in the soprano voice and the 15 positions in which there is not. Indeed, the difference between the average PSLs of melody and accompaniment (ΔMA) was highly correlated with the s.d. ($r=0.78$, $p<0.001$). The largest ΔMA was shown by Cherkassky (10.6 dB), followed by Kentner (9.3 dB) and Hesse-Bukowska (8.6 dB). The smallest difference was exhibited by Horowitz-1972 (0.7 dB), followed by Slenczynska-1975 (1.7 dB) and Bingham (1.8 dB).

However, there was also significant dynamic variation within the melody (and, to a lesser extent, within the accompaniment). Therefore, the standard deviation of the PSLs was also computed for the melody positions separately (ms.d.). The widest melodic dynamic range was shown by Kyriakou (ms.d.=6.6 dB, a range of about 26 dB), followed by Duchâble (6.1 dB) and Ciani (5.7 dB). The narrowest melodic ranges were exhibited by W. Haas (ms.d.=2.0 dB, a range of about 8 dB), Malcuzinsky (2.1 dB), and Horowitz-1972 (2.1 dB). There was a high correlation between the overall and melodic dynamic ranges ($r=0.80$, $p<0.001$), and the multiple correlation of the overall s.d. with both ms.d. and ΔMA was 0.97. However, the correlation between the latter two variables was only 0.31 ($p<0.01$). The dynamic variability thus is best characterized by these two semi-independent measures, which are listed for all performances in Appendix B.

C. The grand average dynamic profile

A grand average dynamic profile was obtained by averaging the dynamic profiles of all 117 performances. This

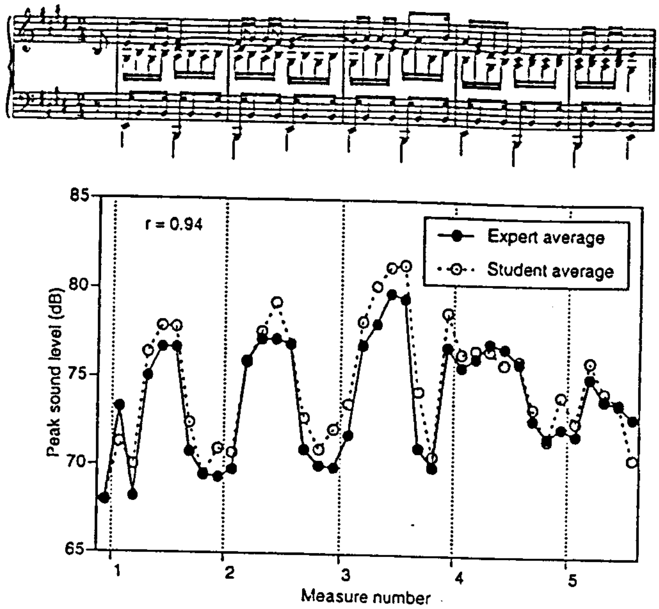

FIG. 1. Grand average dynamic profile of all 117 performances ("expert average"), compared with the grand average profile of multiple performances by 18 advanced students and amateurs ("student average"). A crude score of the music is shown for guidance.

profile is practically identical with the first unrotated principal component (UPC-I) of the data set, except that the latter is computed over standardized dynamic profiles and is expressed likewise in terms of standard scores. The grand average profile ("expert average") is shown in Fig. 1 together with the measured dynamic profile of a synthetic performance on a digital piano, representing the average dynamics of multiple performances of the Chopin excerpt by 18 student and amateur pianists ("student average," see Appendix A). The two profiles have been aligned so that the first data points coincide. They are extremely similar ($r=0.94$, $p<0.001$). One of the few systematic differences is a lower PSL for the students than for the experts in the final position; it is explained by the fact that the students played the excerpt with a final chord, as shown in the crude score above the figure, whereas the experts followed the original score in which the sixteenth-note motion continues in the alto voice. The high similarity of the two average profiles extends an observation made previously about expressive timing to the realm of expressive dynamics, namely that the central tendencies of (advanced) student and expert performances reflect a common performance norm (Repp, 1995b, 1997a, 1997b). With regard to timing, expert performances are distinguished from student and amateur performances by their larger deviations from the common standard, that is, by their often greater individuality. Although a detailed comparison with student performances was not part of the present study, the question about the diversity of dynamic profiles in expert performances was of prime interest.

The grand average profile shows very clearly the difference between melody (peaks) and accompaniment (troughs). The only melody tone that, on average, is as soft as the accompaniment is the initial eighth-note upbeat. The accompaniment stays at a fairly constant dynamic level in bars 1–3 but increases somewhat in intensity in bar 4, which is probably due to the doubling of the accompaniment tones in the right hand (mezzo-soprano plus alto voice). The melody starts out softly, though with a clear dynamic accent on the initial downbeat, and stays at a fairly constant level, about 7 dB above the accompaniment, in bars 1, 2, and 4. The melodic peak in bar 3 (position 5) is marked by a *crescendo* and a dynamic peak which, however, is already reached on the preceding note (position 4). A *decrescendo* occurs in bar 5.

In at least one previous study, a moderate correlation between melodic pitch and intensity has been observed (Repp, 1996a). When this correlation is computed across all primary notes in the present excerpt, it is highly significant ($r=0.82$, $p<0.001$). However, it must be attributed in part to the dynamic contrast between melody and accompaniment rather than to relative pitch height as such. Within the melody, the correlation is smaller but still substantial ($r=0.67$, $p<0.001$). Although this provides statistical support for a relationship between pitch and dynamics, it does not prove a direct causal connection ("the higher the louder"). For example, the general fact that both the melodic pitch contour and the expressive dynamic contour tend to be arched within a phrase may account for the correlation. Closer inspection of the dynamic profile shows a number of local dissociations: For example, in bar 1, intensity (PSL) increases from the first to the second melody tone (positions 1 and 3) even though pitch decreases by a semitone; in bar 3, intensity stays the same between the third and fourth melody tones (positions 4 and 5) even though pitch increases by five semitones; and in bar 4, positions 2 and 3, intensity increases slightly even though pitch decreases by five semitones. Clearly, it would be rash to consider dynamic variation a consequence of pitch variation; at most, pitch is only one of several factors influencing expressive dynamics.

One factor that normally would be expected to affect dynamics is metrical structure. In this very slow music, however, no metrical accents are evident. Such accents would be expected to occur in the first and fifth positions of each bar, but the PSLs in these positions do not differ from those in neighboring positions within the melody or accompaniment. This confirms what musical intuition suggests, namely that the dynamic variation in this piece has only two main functions: to separate the melody from the accompaniment and to give the melody expressive shape. The distinction between these two functions underlies the following analyses of the melody with and without accompaniment.

The grand average dynamic profile is not just a statistical artifact: There are many individual performances whose dynamic profiles are very similar to it. In the principal components analysis (PCA) on the complete dynamic profiles (38 data points), the first unrotated principal component (UPC-I, which is equivalent to the grand average) accounted for 61% of the variance in the data. In the PCA on the melodic dynamic profiles (23 data points), the first unrotated component (mUPC-I) accounted for 48% of the variance. Ninety-six of the 117 performances showed correlations above 0.7 with the UPC-I profile, and 65 showed correlations above 0.7 with the mUPC-I profile. Figure 2(a) and (b) illustrates the individual profiles with the highest respective loadings (correlations). The performances by Kentner and Licad

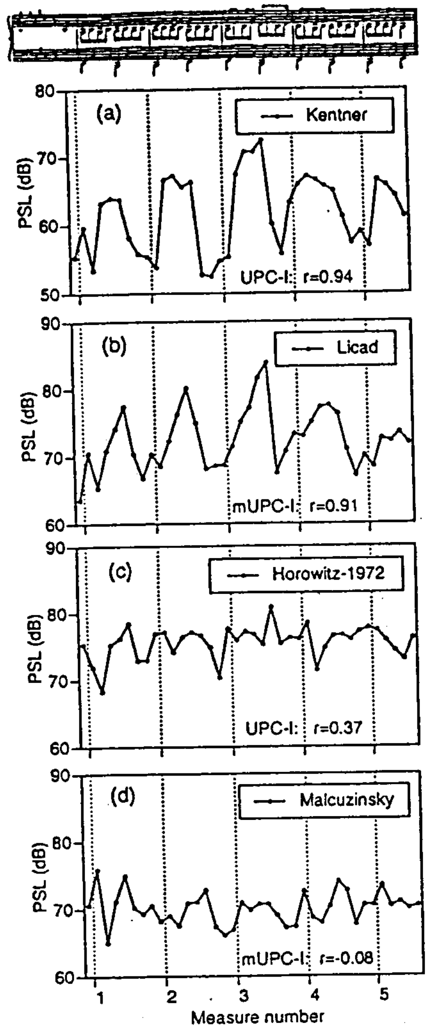

FIG. 2. The performances with (a) the most typical dynamic profile, (b) the most typical melodic dynamic profile, (c) the most atypical dynamic profile, and (d) the most atypical melodic dynamic profile. The relevant PC loadings are shown.

may be considered the ones with the most typical complete and melodic dynamic profiles, respectively. Figure 2(c) and (d) shows the individual profiles with the lowest respective correlations. Horowitz-1972 and Malcuzinsky thus have the least typical dynamic profiles, both of which exhibit very little differentiation between melody and accompaniment. The correlation of Malcuzinsky's melodic profile with mUPC-I is exceptionally low (−0.08); the next-lowest correlation is 0.33. The UPC-I and mUPC-I loadings are moderately correlated ($r=0.66$, $p<0.001$) and may be regarded as indices of typicality. The UPC-I loadings of all performances are listed in Appendix B.

D. Rotated principal components

Although the first principal component (PC) in each PCA accounted for a large percentage of the variance, additional PCs made significant contributions. The traditional criteria for accepting PCs as significant are a discontinuity in the successive percentages of variance accounted for and/or an eigenvalue greater than one (which means that the PC accounts for more than $100/n\%$ of the variance, where n is the number of variables). Neither criterion applied in the present case because no discontinuities were apparent and because the number of data points ($m=38$ or 23) was smaller than the number of variables ($n=117$) and thereby imposed an upper limit on the number of independent PCs. Therefore, the eigenvalue criterion was set at n/m in each analysis; in other words, a PC had to account for more than $100/m\%$ of the variance to be accepted as significant. By this criterion, five PCs were considered significant in each analysis. Together, they accounted for 77% of the variance of the complete profiles and for 75% of the variance of the melodic profiles.

No attempt was made to immediately interpret these additional PCs, each of which accounted for only a small amount of variance. Instead, a Varimax rotation was performed, which distributed the total variance accounted for among the PCs and maximized their range of correlations with the individual profiles. Each rotated PC represents a standardized dynamic profile that is uncorrelated with the other PCs and can be interpreted as a strategy of dynamic shaping. The rotated PCs will be referred to as PC-I through PC-V in the complete profile analysis, and mPC-I through mPC-V in the melodic profile analysis. The five rotated PCs of the complete analysis are related to those of the melodic analysis, but they are not completely congruent. Their numbering follows the output of the statistical program used (SYSTAT) and does not reflect the amount of variance each PC accounts for.

In the complete analysis, PC-I accounts for 17% of the variance. The corresponding melodic profile is mPC-IV, which accounts for 13% of the variance in the melodic profiles. They are in reasonable agreement, as can be seen in Fig. 3(a), and will be referred to collectively as the "type I" profile. The PC-I pattern is characterized by *crescendi* during all melodic gestures, particularly from the initial upbeat to the first downbeat and during the following three-note gesture in bar 1, which leads to an early dynamic peak. The melodic peak in bar 3 is deemphasized relative to the other melodic gestures. A second dynamic peak occurs in bar 4, and there is a small final *crescendo* in bar 5. The accompaniment is generally much softer than the melody, except in the second half of bar 4 where it reaches almost the same level. The mPC-IV pattern differs from the PC-I pattern in that it shows a more pronounced *crescendo* towards the melodic peak in bar 3, but a deemphasis of the gesture-final long notes in bars 1 and 4. No individual profile shows a very high correlation with these patterns. The most representative performances, by Timofeyeva and Richter, respectively, are shown in Fig. 3(b) and (c). Other pianists with relatively high loadings are Haase, Malcuzinsky [see Fig. 2(d)], Koyama, and Vásáry.

Figure 4(a) shows the type II profile, which is defined by two highly correlated PC patterns, PC-II and mPC-V, each of which accounts for 13% of the variance in its respective PCA. The type II profile shows a consistent differentiation of melody and accompaniment and a marked *decrescendo* through bars 4 and 5. Melodic gestures show relatively little dynamic variation, and in particular no *crescendi* (unlike the type I profile). The performance most representative of both PC-II and mPC-V profiles is that by Perlemuter, shown in

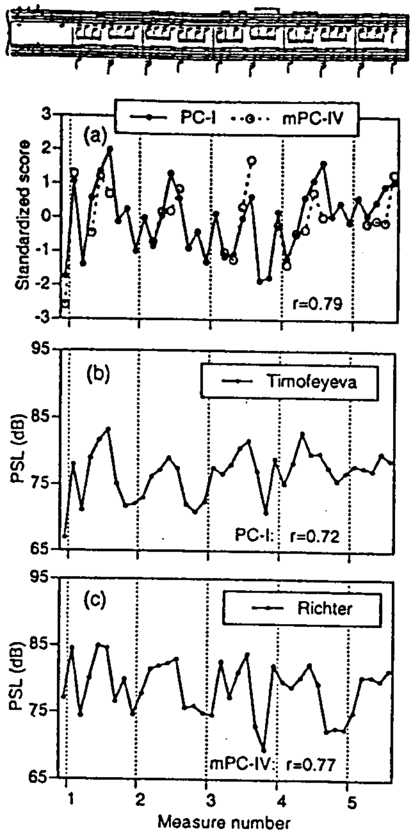

FIG. 3. (a) The type I profile (PC-I) and the corresponding melodic profile (mPC-IV), and the performances with the highest loadings on (b) PC-I and (c) mPC-IV.

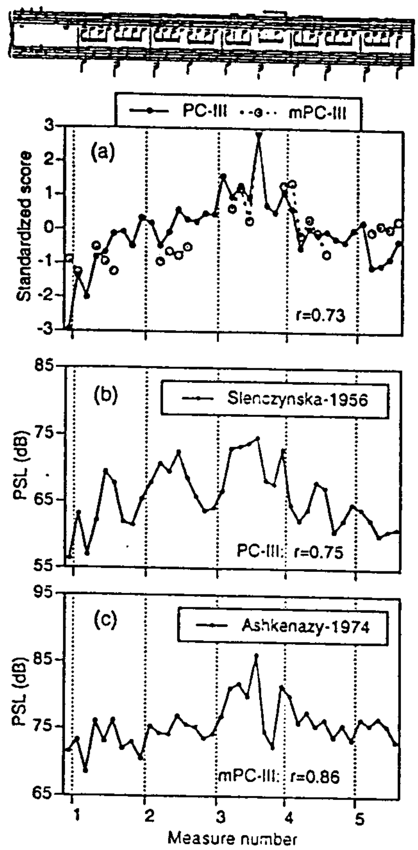

FIG. 5. (a) The type III profile (PC-III) and the corresponding melodic profile (mPC-III), and the performances with the highest loadings on (b) PC-III and (c) mPC-III.

Fig. 4(b). Its melodic gestures are remarkably uninflected. Others with high loadings include Koczalski, Cherkassky, and Goldenweiser.

Figure 5 shows the type III pattern, defined by the moderately correlated PC-III and mPC-III profiles, which account for 13% and 12% of the variance, respectively. It is characterized by a relative lack of differentiation between melody and accompaniment throughout, and by a pronounced dynamic arch that reaches its emphatic peak on the highest melody note. The performance most representative of PC-III, Slenczynska-1956 [Fig. 5(b)], exhibits the dynamic arch but does not show the accent on the melodic peak and maintains a better distinction between melody and accompaniment than the PC-III pattern. The performance of Ashkenazy-1974 [Fig. 5(c)], which has the highest loading by far on mPC-III, does exhibit a lack of differentiation and does show the peak accent but exhibits a less pronounced arch. Other performances that load fairly highly on PC-III include Slenczynska-1975, Bingham, and Donohoe.

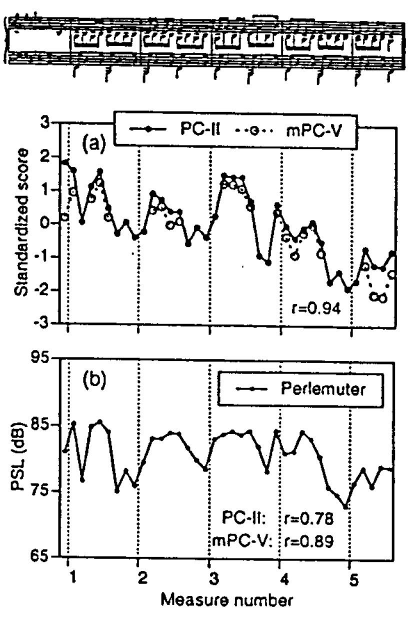

FIG. 4. (a) The type II profile (PC-II) and the corresponding melodic profile (mPC-V), and (b) the performance with the highest loadings on both PC-II and mPC-V.

In the complete profile analysis, PC-IV accounts for the largest percentage of the variance (20%). The same is true for its analog in the melodic analysis, mPC-II (21%). The two patterns are in close agreement, as can be seen in Fig. 6(a), and will be referred to as type IV. This profile is characterized by a very soft beginning in bar 1 but a clear distinction between melody and accompaniment in the following bars. The melodic gestures in bars 3, 4, and 5, however, show an abrupt drop in energy (an effect referred to as *subito piano* in musical parlance) on or before the final long note, which is played as softly as the accompaniment. This is especially noteworthy for the melodic peak in bar 3 which is conspicuously deemphasized, in contrast to the type I and III profiles. The parallel between the soft-spoken three-note melodic gesture in bar 1 and the sudden attenuation of its recur-

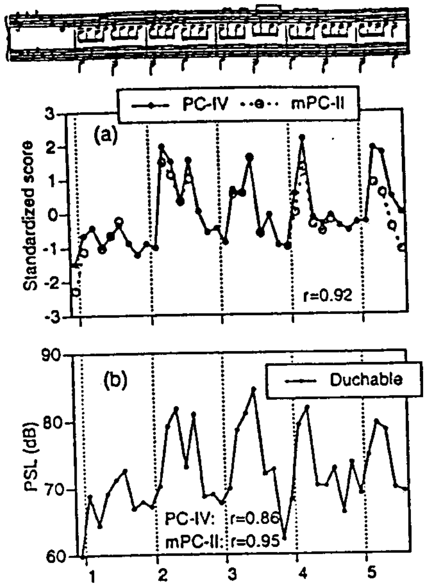

FIG. 6. (a) The type IV profile (PC-IV) and the corresponding melodic profile (mPC-II), and (b) the performance with the highest loadings on both PC-IV and mPC-II.

rence in the second half of the melodic gesture in bar 4 should also be noted. The performance most representative of this interesting strategy, both in its complete and purely melodic versions, is the one by Duchâble, shown in Fig. 6(b). Other performances with high loadings are those by Aide, Hesse-Bukowska, Coop, and Levant.

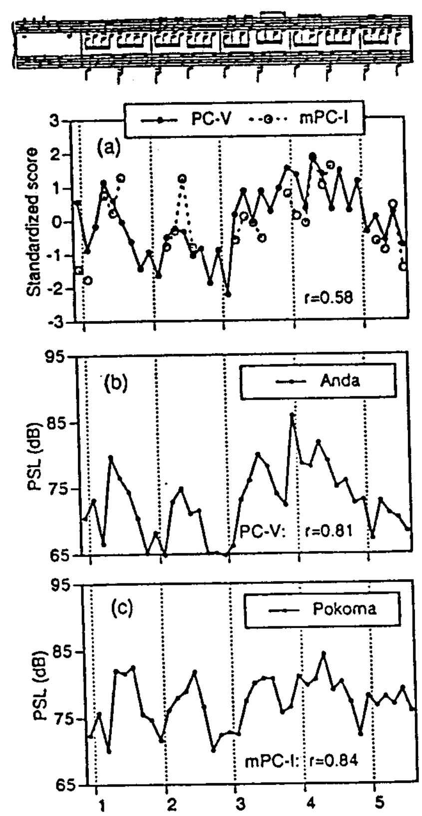

FIG. 7. (a) The type V profile (PC-V) and the corresponding melodic profile (mPC-I), and the performances with the highest loadings on (b) PC-V and (c) mPC-I.

Finally, the type V strategy is illustrated in Fig. 7. Figure 7(a) shows the two defining PC patterns, PC-V and mPC-I, which are only moderately similar. They account for 14% and 15% of the variance in the respective PCAs. PC-V is characterized by a strong initial upbeat, a deemphasis of bar 2, a high dynamic level in bars 3 and 4 with little differentiation of melody and accompaniment, and a final *decrescendo*. The mPC-I pattern deviates in certain details, such as the emphasis on the third note in the four-note melodic gesture of bar 2. The individual performances most representative of these patterns, by Anda and Pokorna, respectively, are shown in Fig. 7(b) and (c). Anda has the highest PC-V loading by far; others with moderately high loadings are Cortot-1942 and Lortie. The mPC-I loadings tend to be higher and include performances by Badura-Skoda and Egorov-1979, among others.

Performances such as those illustrated in the preceding five figures, which load primarily on one of the five PCs, are in the minority. Most performances have modest loadings on two or more of the PCs, which implies that their dynamic profiles can be described as a linear combination of several types. Figures 8 and 9 illustrate four such "hybrid" profiles, derived from the complete PC profiles. Figure 8 shows several linear combinations of PC profiles in which equal weights are given to the different profiles. (The PC profiles were added and then converted back into z-scores, which is equivalent to averaging them.) Figure 9 shows corresponding performance profiles that have approximately equal loadings on several PCs. Thus, Cortot-1933 [Fig. 9(a)] has a type I-V profile [$r=0.74$ with the profile in Fig. 8(a)], Liberace [Fig. 9(b)] has a type II-V profile [$r=0.76$ with the profile in Fig. 8(b)], M. Haas [Fig. 9(c)] has a type III-IV-V profile [$r=0.94$ with the profile in Fig. 8(c)], and Egorov-1978 [Fig. 9(d)] represents a mixture of all five types, though he is leaning towards type IV [$r=0.81$ with the profile in Fig. 8(d)]. It should be noted that the linear combination of all five PCs in Fig. 8(d) is almost identical with the grand average profile (Fig. 1). This follows from the fact that the PCA decomposed the variation among individual profiles into five mutually uncorrelated patterns that account for roughly equal amounts of variance. In terms of these diverse profile types, the most common profile shape (the grand average) seems the most complex, though from another perspective it is the simplest one (viz., the norm). The strategies represented by the rotated PCs are different ways of deviating from the central norm or prototype. The more such ways are adopted simultaneously, the less the net deviation will be.

The pianists' loadings on the different PCs tend to show small negative intercorrelations (computed across all 117 performances), due to the fact that a high loading on one PC implies lower loadings on the others. Some pairs of PCs (I–II, I–III, II–V) exhibit higher negative correlations of their loadings (around -0.3) than other pairs, suggesting that they do not mix as readily as others. However, two-dimensional plots of the PC loadings revealed no clusters or gaps in the distributions, just as in the analysis of the timing data (Repp, 1998a). This suggests that individual performances are fairly uniformly distributed in the multidimen-

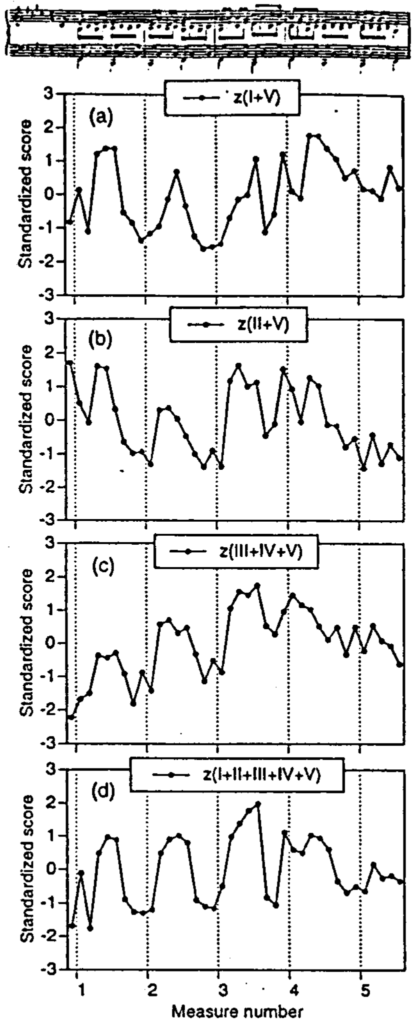

FIG. 8. Four equally weighted combinations of complete PC profiles: (a) PC-I+PC-V, (b) PC-II+PC-V, (c) PC-III+PC-IV+PC-V, and (d) all five PCs.

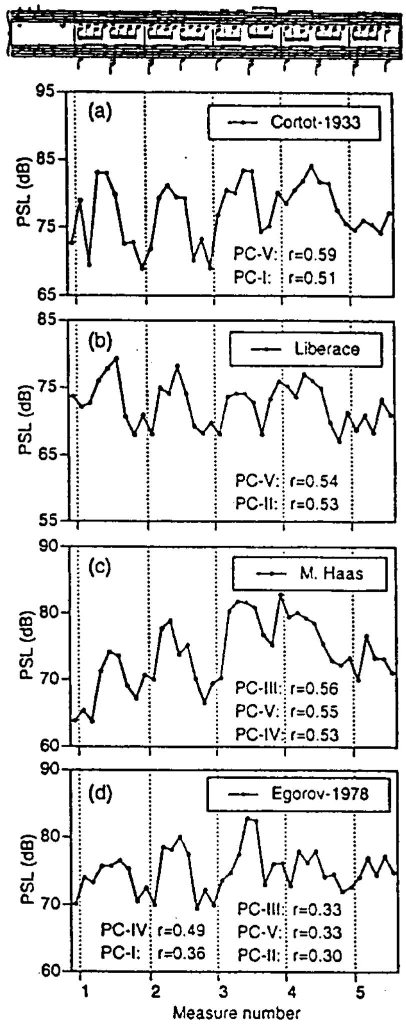

FIG. 9. Performances that best exemplify the combinations shown in Fig. 8.

sional space defined by the PC coordinates. The PC loadings of all performances may be found in Appendix B.

E. Similarity among performances

The intercorrelations among all performance profiles were computed, both for complete and melodic profiles. Each matrix contained 117*116/2=6786 correlations. As expected, two correlations stood out in each matrix—the ones between the two pairs of recordings revealed to be duplicates in the timing analysis (Repp, 1998a). The two versions of Zarankin, which derived from different CDs (one was misattributed to a pianist named van der Voss), were perfectly correlated. The two versions of Aide (copied twice from the same recording by mistake) correlated 0.990 (0.981 for the melody only). These correlations were much higher than the next-highest correlations, which proves beyond any doubt that these pairs of performances were in fact identical. The imperfect correlations for Aide must have been due to some kind of small measurement error.

The highest correlation between the complete dynamic profiles of two different performances was 0.929 (Kahn and Kentner). This is about as similar as dynamic profiles of different artists get. Figure 10(a) shows them superimposed. [Kentner's profile is also shown in Fig. 2(a).] Both are examples of typical profiles that correlate highly with the grand average profile and load on several rotated PCs. The highest correlation between the melodic profiles of different performances was 0.924 (Slenczynska-1975 and Renard), as shown in Fig. 10(b). They are both type III profiles [cf. Fig. 5(a)]. There were only four negative correlations among the complete profiles, the lowest of which was −0.109 (Slenczynska-1975 and Malcuzinsky). The contrast between Slenczynska's highly arched profile and Malcuzinsky's flat profile may be gauged by comparing Figs. 10(b) and 2(d). Negative correlations were more frequent among the melodic profiles and reached a minimum of −0.508 (Haase and Crown). Figure 10(c) shows that Haase makes strong *crescendi* during melodic gestures and especially in bar 5 (a type I profile), whereas Crown tends to deemphasize melodic peaks and makes a *decrescendo* at the end (a type II profile).

Seven pianists were represented by two different performances each, recorded many years apart except in the case of Egorov, whose two recordings are separated by only one year. However, none of these artists showed exceptionally high correlations between his or her two dynamic profiles. The highest correlations were shown by Cziffra (0.867 com-

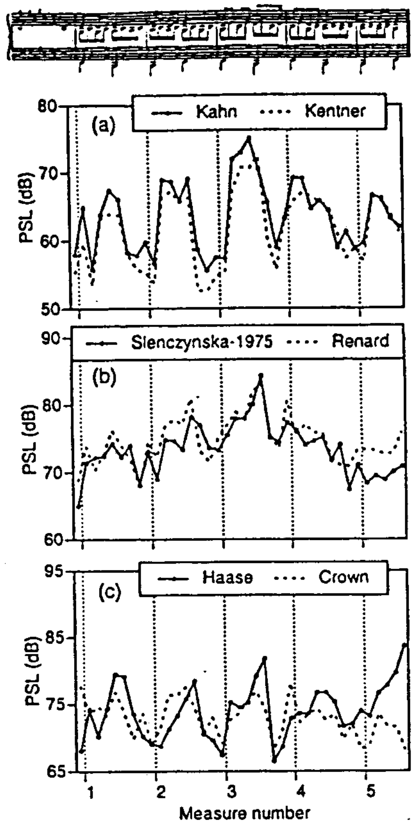

FIG. 10. The two most similar performances in terms of (a) complete dynamic profiles and (b) melodic dynamic profiles; also, (c) the two most dissimilar performances in terms of complete profiles.

plete, 0.758 melodic), Slenczynska (0.728 complete, 0.834 melodic), and Cortot (0.777 complete, 0.680 melodic). Arrau, Ashkenazy, and Egorov showed only moderate correlations, the last a very low melodic correlation. The lowest consistency was exhibited by Horowitz (0.270 complete, 0.224 melodic). Thus there was little evidence for the maintenance of an individual dynamic profile across many years, or even across one year in the case of Egorov. In general, little is known about pianists' consistency in this respect, though an individual dynamic profile can certainly be replicated in repeated takes on the same day (Repp, 1996a).

F. Sociocultural variables

As described in Repp (1998a), information about five sociocultural variables was available: artists' gender, nationality (i.e., country of birth, without regard to educational history), birth date, recording date, and age at the time of recording (which was negatively correlated with birth date but unrelated to recording date). Except for gender, the data were somewhat incomplete but sufficient for a preliminary investigation. The dependent variables were the two indices of dynamic range (ms.d., the standard deviation of the melodic PSLs, and ΔMA, the difference between the average PSLs of melody and accompaniment), the two typicality indices (UPC-I and mUPC-I loadings), and the rotated PC loadings in the PCAs on the complete and melodic profiles.

Recording dates were unrelated to measures of dynamic range (which suggests, surprisingly, that old recordings were not of more limited dynamic range than new CDs) but marginally related to typicality: The correlation with mUPC-I loadings reached significance ($r=0.25$, $p<0.05$); that with UPC-I loadings did not. Thus, there was a slight tendency for more recent recordings to exhibit a more typical melodic dynamic profile, which is consistent with the frequent claim of music critics that performances are less individual now than many years ago. Recording date showed some more substantial correlations with dynamic strategies. In particular, it was negatively correlated with the type II pattern ($r=-0.40$, $p<0.001$, with PC-II loadings; $r=-0.33$, $p<0.001$, with mPC-V loadings). This pattern, which is characterized by a relatively uninflected melody and a marked drop in dynamic level towards the end of the phrase (see Fig. 4), is thus more frequently encountered in older recordings. Weaker, positive correlations of recording date with the type I pattern ($r=0.23$, $p<0.05$, with PC-I loadings; $r=0.32$, $p<0.01$, with mPC-IV loadings) and with the type V pattern ($r=0.25$, $p<0.05$, with PC-V loadings; $r=0.31$, $p<0.01$, with mPC-I loadings) indicated that these strategies are encountered more often in recent recordings.

Artists' birth dates, like recording dates, showed a slight positive correlation with mUPC-I ($r=0.22$, $p<0.05$), a negative correlation with type II ($r=-0.33$, $p<0.001$, with PC-II loadings; $r=-0.31$, $p<0.01$, with mPC-V loadings), and a marginal positive correlation with type I ($r=0.22$, $p<0.05$, with PC-I loadings), but a stronger positive correlation with type V ($r=0.31$, $p<0.01$, with PC-V loadings; $r=0.45$, $p<0.001$, with mPC-I loadings). Thus, a melodic profile that gives special emphasis to bar 4 but deemphasizes bar 2 (type V) is favored more by younger generations of pianists.

Age at the time of recording was related weakly to two melodic dynamic strategies: positively to type II ($r=0.24$, $p<0.05$, with mPC-V loadings) and negatively to type V ($r=-0.26$, $p<0.01$, with mPC-I loadings). These correlations have the opposite sign of those for birth date, which reflects the negative correlation between birth date and age at the time of recording. Thus, the type II pattern is more associated with older artists, the type V pattern more with younger artists.

Effects of nationality were difficult to assess because of the many countries represented. As in the earlier timing analysis, the analysis was restricted to three well-represented nationalities with strong traditions of Chopin performance: French, Polish, and Russian. However, one-way analyses of variance (ANOVAs) revealed no significant effects of this independent variable on any of the dependent variables.

Gender was similarly analyzed in one-way ANOVAs, but making use of the complete sample. One significant difference emerged: Female pianists exhibited a significantly greater dynamic range in the melody [$F(1,113)=7.1$, $p<0.01$]. The average ms.d. was 3.9 dB for women as compared to 3.4 dB for men. This greater dynamic inflection may indicate a more liberal display of emotion by female than by male artists. A second effect fell just short of significance: Women exhibited somewhat higher average mUPC-I loadings (0.729) than men (0.658) [$F(1,113)=3.66$, $p<0.06$], which means that the female pianists in the sample tended to produce somewhat more typical or conventional melodic dynamic patterns than did the male pianists.

G. Relationships with expressive timing

Todd (1992) has described a tendency for dynamics to covary with timing, which may be summarized as "the faster, the louder" (or vice versa). If timing is measured in terms of IOI duration, this implies a negative correlation. While there are numerous exceptions to this rule, it may be considered as a default or "unmarked" case whose violation is "marked" and thereby conveys special expressive effects. It is not clear, however, whether this reasoning can be applied at the local level considered here.

To begin with, the grand average timing profile (Repp. 1998a, Fig. 4) was compared with the grand average dynamic profile (Fig. 1). Their overall correlation was positive ($r=0.44$, $p<0.01$), contrary to Todd's rule.[2] The reason for this is that the accompaniment during sustained melody notes is generally played both faster and softer than the melody. A more meaningful correlation, therefore, is that for the melody notes only. It was negative but failed to reach significance ($r=-0.33$). Moreover, it dropped to -0.05 when the very long IOI and the corresponding PSL of the first downbeat were omitted. Thus, there was not really any relation between typical timing and typical dynamics within the melody. The same was true within the accompaniment ($r=0.08$). There was no evidence for any nonlinear relationship either.

Next, the correlations between the grand average timing profile and the melodic dynamic PC patterns were examined. There was a sizeable positive correlation with mPC-IV ($r=0.64$, $p<0.001$), the dynamic type I pattern. It will be recalled that the type I profile is characterized by *crescendi* within melodic gestures, and these *crescendi* correspond to the *ritardandi* that typically occur as well. Intuitively, however, type I does not seem "marked," even though it violates Todd's rule. The *subito piano* strategy of type IV seems much more deserving of that epithet.

Next, the correlations between the grand average melodic dynamic profile and the four timing PC profiles were computed. The only significant correlation was with timing PC-III ($r=-0.48$, $p<0.05$), which is characterized by an exceptionally long first (initial downbeat) IOI. With that IOI omitted, the correlation withered.

Finally, the $4\times5=20$ correlations between the PC profiles for melodic timing and the PC profiles for melodic dynamics were examined. The only sizeable correlation was between the PC-I for timing and the dynamic mPC-IV ($r=0.67$, $p<0.001$). This relationship is essentially the same as the one between the grand average timing profile and mPC-IV, for the PC-I for timing is similar to the grand average timing profile.

Even though there were no relationships between major timing strategies and dynamic strategies, some interdependence of the two expressive parameters at the individual level seemed possible. Therefore, the correlations between melodic timing and dynamics were computed for all 115 individual performances. Only 18 correlations reached significance ($|r|>0.43$, $p<0.05$), and all but one were negative. The highest negative correlations were shown by Fou Ts'ong ($r=-0.74$, $p<0.001$) and Karolyi ($r=-0.67$, $p<0.001$), whereas the only significant positive correlation was exhibited by Malcuzinsky ($r=0.57$, $p<0.01$). It may be recalled that Malcuzinsky has the most atypical melodic dynamic profile [Fig. 2(d)]. While it is possible that timing and dynamics were functionally linked for these 18 artists, it seems just as likely that they pursued independent strategies that just happened to show some correspondence. Moreover, the initial downbeat probably accounted for the negativity of most correlations (see above).

To examine whether there was any relationship between the relative typicality of individual timing profiles and dynamic profiles, the correlation between the respective UPC-I loadings was computed across all performances. It was zero. Next, possible associations of the PC patterns for timing and dynamics were investigated by computing the $4\times5=20$ cross correlations between the respective PC loadings. While some of these correlations reached significance due to the large degrees of freedom, they were too small ($r<0.3$) to have any explanatory value.

A final possibility examined was that the degrees of timing modulation and of dynamic modulation might be associated with each other, regardless of the pattern exhibited. However, the relevant correlations were nonsignificant.

In summary, then, these analyses suggest that expressive timing and expressive dynamics are independently controlled at the within-phrase level of detail considered here. The possibility remains that there are even more local dependencies than were considered in these analyses, such as within melodic gestures. However, this would be difficult to prove, given the small number of data points.

III. GENERAL DISCUSSION

This is the first study of individual differences in expressive dynamics in a large sample of expert performances. Individual patterns of expressive dynamics proved to be at least as diverse as individual patterns of expressive timing. For timing, four "strategies" had been identified, one of which concerned only the initial downbeat IOI (Repp, 1998a). For dynamics, there were five strategies that affected the dynamic shaping throughout the phrase. Moreover, for timing there was a dominant strategy (PC-I) which resembled the grand average timing profile, whereas for dynamics none of the five strategies was particularly prevalent or similar to the grand average profile. In a sense, therefore, expressive dynamics offers more opportunities for the exhibition of individuality than does expressive timing, at least in the musical excerpt studied. However, because timing tends more towards a central norm and because there are fewer categorically distinct options of deviating from this norm, individuality in timing may be more conspicuous when it occurs.

The central norm is identified with the grand average profile. This seems to be as valid for dynamics as it is for timing. For both expressive dimensions, the majority of performances shows a moderate to high degree of similarity with the grand average profile. This profile is a statistical summary of all the individual variation, and as such it reveals the *potentialities* of timing or dynamics. An appropriate metaphor may be to regard a musical passage as a flexible ribbon varying continuously in thickness or stiffness, due to its internal structure. Points of high flexibility are those where the music is less cohesive and invites the performer to "stretch" or "bend" it. A typical (mainstream, conventional, textbook) performance tries to realize all these potentialities simultaneously, as would a sophisticated computer algorithm (yet to be devised) that generates expressive variation deterministically from a complete structural representation. An individual (original, creative, unconventional) performance deviates from the normative pattern, following one or more of several possible expressive strategies. Such strategies necessarily highlight some aspects of the musical structure while deemphasizing others, but it is far from clear that this is the artist's purpose in expressive performance. Some strategies create contrast between notationally similar structures by giving them different temporal or dynamic shapes.

Dynamic strategies seem less constrained by the musical structure than timing strategies. Timing within a rhythmic group usually has to exhibit continuity and smoothness, or else the listener will perceive momentary hesitations and accelerations that are aesthetically undesirable. This restricts timing to basic acceleration–deceleration shapes (Todd, 1992). Expressive dynamics does not seem to be restricted in this way. Besides smooth *crescendi* and *decrescendi*, a variety of other accent patterns is possible, both in conformity with and against the underlying metrical grid. Whereas timing imparts a particular movement to a musical gesture, gives it coherence, and separates it from other gestures, dynamics give it a particular "flavor" or character. Short of wild exaggeration of accents or contrasts, dynamic patterns do not seem to sound deviant or irregular as easily as do timing patterns. These observations are rather speculative, however, and need further investigation.

Many performers blend different strategies into novel combinations. While this may seem especially creative, it is an interesting paradox that, the more strategies are combined, the more the resulting pattern resembles the grand average profile. The central norm is the combination of all possible strategies of deviating from it; it is synthetic. A truly original performance follows a single deviant strategy; in this sense it is analytic.[3] There is another paradox here, however, in that the PCs—the result of a Varimax rotation—were determined automatically in order to maximize the number of performances with high loadings. As a result, for every truly original performance (i.e., one loading highly on a single PC) there are several other performances that are original in more or less the same way. Where, then, does true originality lie? The PCs could have been rotated to a different configuration, one that does not satisfy the Varimax criterion but accounts for just as much variance in the data. Would the PCs of any such rotation be associated with *uniquely* individual performances? This seems unlikely in view of the relatively uniform distribution of the PC loadings in the two-dimensional spaces spanned by pairs of PCs. It seems likely that any rotated PC would be associated with small groups of performances. Thus, originality can only be equated with nontypicality, i.e., distance from the central prototype, not with uniqueness or with a distinct region in the space of possibilities. Of course, the large number of degrees of freedom available in expression makes any individual performance effectively unique, but there are always other performances that are similar to it in one way or another. In other words, the territory of expressive possibilities has been well explored by artists, and while there are many uninhabited regions, they probably will remain so because they are inhospitable to both performers and listeners.

As was already pointed out in the earlier article on timing (Repp, 1998a), the psychological status of the expressive strategies remains unclear. At this point, they are merely descriptive coordinates that make it possible to talk about the differences among 115 performances without having to describe each of these performances. It does not seem to be the case that different expressive strategies reflect different structural interpretations of the music, at least not in a categorical sense; nor do "pure" strategies seem to be more pure in any cognitive sense than "mixed" strategies. While structural interpretation draws on a limited set of categorically distinct possibilities, expressive shaping draws on a circumscribed but continuous and therefore effectively unlimited range of possibilities. The musical excerpt studied here does not really contain any structural ambiguities, at least not any that performers feel compelled to resolve, and therefore the performance analyses do not provide any evidence of categorically different structural interpretations. Rather, as already argued in Repp (1998a), they reveal continuous individual variability within constraints set by a single structural frame.

The apparent independence of timing and dynamics in the Chopin excerpt studied here may seem surprising. An interdependence between these two expressive dimensions may well be found in other kinds of music that have a more pronounced rhythmic structure (e.g., various dance forms). The present negative result may be due to the slow, almost arhythmic nature of the music, which offers maximal freedom in expressive shaping and was selected for that reason. Certainly, the independence of timing and dynamics increases the performer's degrees of freedom tremendously. Then there are additional factors that were not even considered here because they are too difficult to analyze objectively in acoustic recordings (such as "touch," articulation, and pedaling). And it should not be forgotten that timing and dynamics each have three independent aspects: basic tempo and dynamic level, timing modulation and dynamic modulation, and timing pattern and dynamic pattern. Basic tempo and dynamic level primarily set the emotional tone of a performance: calm versus excited and gentle versus forceful, respectively. Degree of modulation defines a nonspecific dimension ranging from understatement to exaggeration. Patterns give shape and character. Expressive timing patterns are forms of movement; they govern the variable rate at which the musical sound structure unfolds. Dynamic patterns are part of the sound structure itself. They are a part of what is unfolding, whereas timing governs *how* this unfolding is taking place.

It is sometimes thought that music does not have any expressive dynamics, because the intensities at which individual notes are to be played are not specified in the score—they are "added" by the performer. However, this is a serious misconception. Music in which all notes are of equal intensity does not sound right, especially when distinctions among voices are to be made. It is not the music represented in the score but a distortion of it. Therefore, music must always have *some* expressive dynamics, just as it must have *some* expressive timing (unless the music is specifically intended to be performed without expression). Elsewhere, the author has argued that a typical timing pattern is the most appropriate default timing pattern for a piece of music (Repp, 1995b, 1997a). This argument may now be extended to the typical dynamic pattern.[4] A performance possessing both of these patterns (plus appropriate articulation, pedaling, etc.) will be a perfectly acceptable realization of the score, and may in fact be considered as being the best approximation to what the score represents. Individual performances may deviate from these normative expressive patterns, but they do not "deviate from the score."

The present research also extends to dynamics a finding obtained repeatedly for timing (Repp, 1995b, 1997a, 1997b): The grand average profile of a group of advanced student pianists and amateurs is extremely similar to that of a large sample of expert recording artists. This suggests that the same norm of expressive shaping underlies performances at different levels of expertise. However, expert pianists' performances are much more diverse in timing than student performances, which tend to stay close to the norm. Without any doubt, this is also true for dynamics, even though no detailed analysis of individual differences in students' dynamic profiles has yet been conducted for the Chopin excerpt. Repp (1996a) commented on the narrow range of variation of student pianists' dynamic profiles in Schumann's "Träumerei." This relative conservatism of student pianists is not at all surprising and may have a number of reasons, among them relative inexperience, lesser creativity compared to famous artists, and the increasing homogenization of classical music performance due to the influence of recordings and the competition circuit.

In this article and its predecessor (Repp, 1998a), evaluative comments have been studiously avoided; the analyses have been purely descriptive and objective. Yet, the reader may have been wondering about the relation of all this expressive diversity to the perceived quality of the performances. This topic will be addressed in the third and final installment in this series of papers (Repp, in preparation).

ACKNOWLEDGMENTS

This research was supported in part by NIH Grant No. MH-51230. The principal measurements were performed during an unfunded period during which the author's research was made possible only by the love and support of his wife, which is gratefully acknowledged here. Thanks are also due to Paul Buechler for assistance with the measurements reported in Appendix A, and to Nigel Nettheim and an anonymous reviewer for helpful comments on the manuscript.

APPENDIX A: PRELIMINARY ACOUSTIC MEASUREMENTS

To investigate how the overall intensities of piano sonorities depend on the relative intensities of their component tones, acoustic measurements were performed on two performances whose individual tone intensities were known.

1. Method

In the course of several previous studies, 18 advanced student and amateur pianists had played the Chopin Etude excerpt repeatedly (three or ten times) on a Roland RD-250s digital piano and had been recorded in MIDI format. These MIDI data were imported into a spreadsheet program, and the key-press velocities of individual notes were averaged across each pianist's repeated performances and then across all pianists. These grand average velocities, assumed to represent a typical dynamic pattern for the music, were combined with MIDI instructions specifying constant sixteenth-note inter-onset intervals of 500 ms as well as synchronous note onsets and offsets (following the score literally), without pedal.[5] This performance was played under computer control on the digital piano. It will be referred to as the SA (student average) performance.

A second performance was played on the digital piano by a skilled young pianist (H.S.) and recorded in MIDI format. It differed from the computer-generated performance in that it was expressively timed and included both pedaling and small asynchronies among nominally simultaneous note onsets. In particular, the melody notes tended to precede the notes in the other voices, especially those played by the same (right) hand, as is commonly found in expressive performance (Palmer, 1996b; Repp, 1996b). This will be referred to as the HS performance.

Two reduced versions were created from the MIDI instructions for each of these two performances. In one, the MIDI velocities of all secondary notes were set to zero, so that only the primary notes remained. In the other, the primary notes were moreover shortened to 100-ms duration and, in the HS performance, the pedal instructions were deleted. This eliminated all acoustic overlaps among successive primary tones caused by sustained melody tones, decay of immediately preceding damped tones (see Repp, 1995a), and pedaling.

All six versions were played back under computer control on the digital piano, were recorded electronically onto digital audio tape, and then were resampled by a Macintosh Quadra 660AV computer at a rate of 22.055 kHz. Using SIGNALYZE software, the root-mean-square amplitude envelopes of the digitized waveforms were computed with a sliding rectangular integration window of 30 ms, and subsequently an automatic peak-picking routine was employed to determine the peak amplitude following each note onset. These peak amplitudes were then converted into peak sound levels (PSLs) in decibels (dB).

2. Results and discussion

Figure A1 compares the dynamic profiles for the full performances with the versions containing primary notes

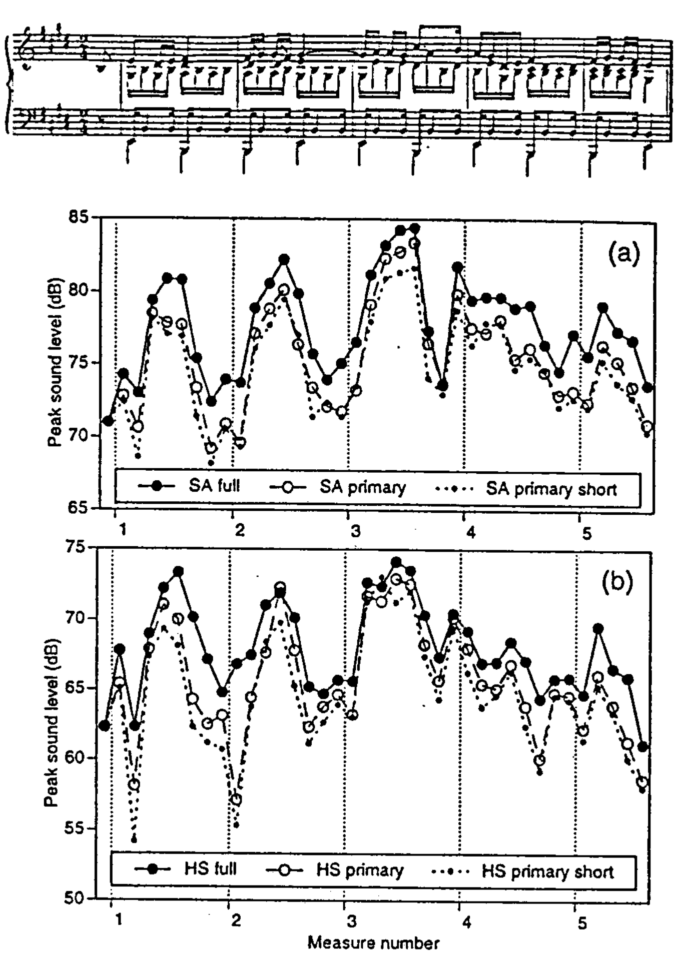

FIG. A1. Measured peak sound levels of three versions of (a) the computer-generated SA performance and (b) the human HS performance. The three versions are: full (all notes), primary notes only, and short (truncated) primary notes.

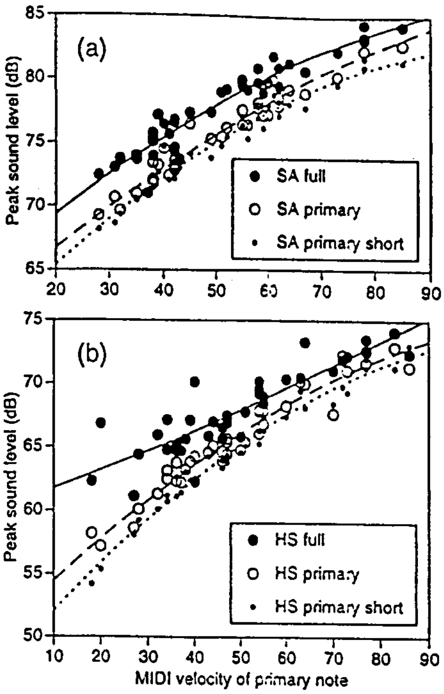

FIG. A2. Measured peak sound levels as a function of the MIDI velocity of the primary notes for the three versions of (a) the SA performance and (b) the HS performance.

only. (The difference in average absolute sound level between the SA and HS performances is meaningless.) The PSLs of the initial upbeat are identical in the three versions of each performance because it was not accompanied by any other tone. As expected, the PSL values for the full performance are generally higher than those for the primary tones only, due to the contribution of the secondary (lower-pitched) tones and the sustained melody tones. The PSL values for the primary tones in turn are somewhat higher than those of the short primary tones, reflecting the contribution of the overlaps of successive tones. In the case of the SA performances [Fig. A1(a)], the dynamic profiles of the three versions are highly similar: The full and primary profiles correlate 0.96, and the primary and primary-short profiles correlate 0.98. Even if only the 23 melody-note positions are considered, the correlations are still of the same magnitude. For the HS performances [Fig. A1(b)], the correlation between the full and primary profiles is not quite as high ($r = 0.89$), whereas that between the primary and primary-short profiles is 0.97. If only the melody notes are considered, the corresponding correlations are 0.95 and 0.97, respectively. It can be seen in Fig. A1(b) that the major discrepancies between primary and full PSLs occur during the accompaniment passages, where the primary notes are relatively soft.

This correlational evidence suggests that the overall dynamic profile is quite representative of the dynamic profile of the primary notes, especially of the melody notes. The analysis went one step further by investigating whether the overall PSLs could be predicted from the known MIDI (key-press) velocities of the individual notes. Figure A2 shows the measured PSLs as a function of the MIDI velocities of the primary notes. Quadratic functions have been fitted to the data points. For the primary and primary-short versions, these functions are very similar to the function obtained in a previous study, where PSL measurements were averaged across many different pitches on the same digital piano (Repp, 1993; see Repp, 1997c, Fig. 1). Deviations of individual data points from the function may reflect imperfect calibration of different pitches on the electronic instrument and/or measurement error of some kind. The function for the full SA performance is also similar, even though many additional notes contributed to the overall PSLs. Only the function for the full HS performance is different and reflects larger contributions of the additional notes when the MIDI velocity of the primary note is low than when it is high.

Linear regression analyses were subsequently conducted, with the full performance PSLs as the dependent variable and the MIDI velocities of the individual notes as the independent variables. The velocities of the primary notes constituted the first independent variable, and the velocities of additional mezzo, alto, tenor, and bass notes yielded four additional independent variables that were padded with zeroes in positions where there were no note onsets.

A standard regression analysis on the SA full performance data accounted for 93% of the variance ($R = 0.97$). Naturally, the largest contribution was made by the primary-note velocities, which alone accounted for 89% of the variance, but small additional contributions were made by the secondary notes in each voice. In a subsequent stepwise re-

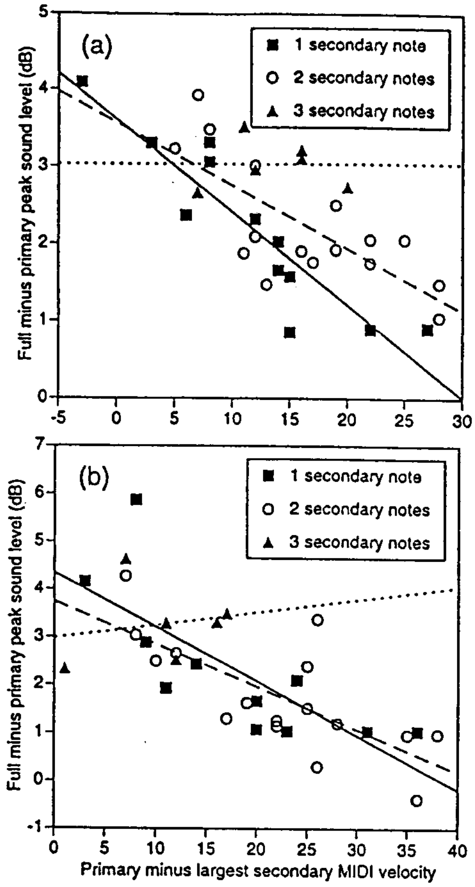

FIG. A3. The increase in peak sound level due to secondary notes as a function of the velocity difference between the primary note and the loudest secondary note, and of the number of secondary notes, for (a) the SA performances and (b) the HS performances.

gression analysis, however, only the contribution of secondary alto notes reached significance.

A standard regression analysis on the HS full performance data accounted for 85% of the variance ($R=0.92$). The primary-note velocities alone accounted for 78% of the variance. In a stepwise regression analysis, the contributions of secondary mezzo, tenor, and bass notes reached significance, but, in contrast to the SA performance results, secondary alto notes made no contribution at all. This may have been due to a tendency of secondary alto notes onsets to lag behind primary melody note onsets. The variance not accounted for must be attributed to varying asynchronies and tone overlaps. No attempt was made to enter these additional variables into a regression analysis, though they would have to be included in a complete predictive model.

Figure A3 illustrates graphically the contribution of secondary notes. The difference between the full and primary PSLs (i.e., the increase in PSL due to secondary notes) is shown as a function of the MIDI velocity difference between the primary note and the most prominent secondary note in the same position, regardless of voice. The solid regression line represents positions in the music where there was only a single secondary note. Clearly, this extra tone contributed to the overall PSL in proportion to its intensity relative to the primary tone.[6] The dashed regression line represents positions in which there were two secondary notes accompanying a primary note. The additional contribution of the second (softer) secondary note is reflected in the difference between the solid and dashed regression lines; such a difference is evident only for the SA performances [Fig. A3(a)]. The dotted regression line represents the five positions in the music where there were three secondary notes; here there is no longer any relationship between the relative velocity of the strongest of them and the PSL increase, but some contribution of the third secondary note is suggested, at least for the SA performances.

Regression analyses were conducted with the independent variables recoded as the velocities of primary, second, third, and fourth notes, ordered according to velocity magnitude and regardless of voice. The standard regression equations accounted for the same amounts of variance as previously. In stepwise regression analyses, primary, second, and third notes made significant contributions in the SA performances; primary, second, and fourth notes in the HS performances.

The y-axis intercepts of the regression lines in Fig. A3 indicate that addition of an equally intense tone of lower pitch to a primary tone results in an increase of about 4 dB in overall PSL. If peak amplitudes were strictly additivite, the expected increase would be 6 dB. Palmer and Brown (1991) have presented data showing that the peak amplitudes of two simultaneous piano tones of different pitch are nearly additive: The combined amplitude increased by 81% to 96%, which translates into 5.2 to 5.8 dB. The smaller increases observed here may be due to smoothing by the 30-ms temporal window used in determining the PSLs. The similar y-axis intercepts of the regression lines for the AS and HS performances suggest that onset asynchronies in the HS performances were not responsible.

APPENDIX B: PERFORMANCE STATISTICS

ms.d.=standard deviation of melodic PSLs (dB).

ΔMA=difference between average PSLs of melody and accompaniment (dB).

UPC-I=loadings on first unrotated PC for the complete profile (index of typicality)

PC-I through PC-V=loadings on rotated PCs for the complete profile

Pianist	ms.d.	ΔMA	UPC-I	PC-I	PC-II	PC-III	PC-IV	PC-V
Aide	3.1	4.6	0.810	0.201	0.264	0.137	0.806	0.298
Aide 2	2.9	4.9	0.797	0.161	0.325	0.083	0.800	0.314
Anda	4.3	6.6	0.740	0.210	0.290	0.177	0.204	0.811
Anievas	3.6	5.7	0.743	0.118	0.249	0.289	0.650	0.301
Arrau 1930	3.9	6.2	0.837	0.142	0.457	0.506	0.529	0.251

Pianist	ms.d.	ΔMA	UPC-I	PC-I	PC-II	PC-III	PC-IV	PC-V
Arrau 1956	4.5	4.6	0.798	0.458	0.153	0.379	0.669	0.038
Ashkenazy 1959	3.4	5.5	0.787	0.227	0.368	0.254	0.521	0.366
Ashkenazy 1974	3.4	3.2	0.721	0.168	0.262	0.692	0.237	0.324
Backhaus	3.1	6.0	0.857	0.552	0.353	0.315	0.442	0.230
Badura-Skoda	3.3	5.4	0.738	0.565	0.024	0.081	0.319	0.608
Berezovsky	3.7	4.1	0.811	0.570	0.047	0.461	0.268	0.466
Bingham	3.9	1.8	0.672	0.503	0.071	0.734	0.172	0.061
Binns	3.0	3.4	0.695	0.524	0.260	0.375	0.179	0.238
Biret	2.9	2.6	0.762	0.253	−0.111	0.460	0.485	0.575
Brailowsky	3.8	5.8	0.737	0.251	0.648	0.350	0.337	0.107
Browning	2.4	2.7	0.792	0.522	0.407	0.555	0.130	0.228
Cherkassky	5.4	10.6	0.850	0.299	0.679	0.340	0.216	0.447
Ciani	5.7	7.1	0.809	0.240	0.371	0.578	0.413	0.235
Ciccolini	3.4	6.0	0.802	0.654	0.158	0.140	0.376	0.414
Cliburn	3.8	3.6	0.769	0.544	0.081	0.417	0.492	0.129
Coop	4.5	7.2	0.835	0.233	0.142	0.261	0.752	0.388
Cortot 1933	3.0	6.2	0.874	0.506	0.383	0.252	0.291	0.538
Cortot 1942	3.8	4.9	0.799	0.291	0.367	0.367	0.141	0.693
Costa	4.1	7.7	0.806	0.310	0.582	0.312	0.203	0.464
Crown	2.3	3.5	0.513	0.148	0.760	−0.089	0.199	0.170
Cziffra 1954	2.9	6.3	0.903	0.480	0.189	0.339	0.504	0.466
Cziffra 1981	3.7	6.5	0.841	0.497	0.052	0.182	0.674	0.364
Darré	5.3	6.6	0.663	−0.004	0.600	0.195	0.298	0.449
Donohoe	4.1	2.8	0.676	0.055	0.493	0.725	0.240	0.095
Drzewiecki	4.5	5.2	0.828	0.232	0.102	0.586	0.483	0.442
Duchâble	6.1	5.8	0.687	0.122	0.169	0.277	0.858	−0.008
Egorov 1978	2.8	4.0	0.818	0.357	0.300	0.328	0.493	0.326
Egorov 1979	2.4	5.8	0.802	0.604	0.267	0.079	0.320	0.493
Ellegaard	4.3	6.0	0.733	0.016	0.519	0.240	0.638	0.201
Entremont	3.0	2.5	0.602	0.555	0.263	0.346	0.072	0.146
Farrell	3.0	5.1	0.720	0.620	0.201	−0.001	0.417	0.298
Fou Ts'ong	5.0	6.5	0.836	0.201	0.082	0.481	0.599	0.465
François	3.6	4.6	0.799	0.295	0.528	0.431	0.414	0.144
Goldenweiser	3.5	8.2	0.770	0.151	0.679	0.180	0.308	0.458
Goldsand	3.2	6.4	0.844	0.610	0.398	0.232	0.507	0.092
Goodman	2.5	6.0	0.791	0.252	0.396	0.160	0.526	0.403
Haas, M.	5.0	5.5	0.819	0.053	0.146	0.556	0.529	0.546
Haas, W.	2.0	2.8	0.711	0.439	0.333	0.228	0.097	0.540
Haase	3.5	5.4	0.705	0.712	0.151	0.167	0.326	0.169
Harasiewicz	3.6	6.4	0.861	0.313	0.336	0.246	0.408	0.622
Hesse-Bukowska	4.5	8.6	0.883	0.422	0.266	0.227	0.775	0.182
Hobson	2.5	4.2	0.782	0.589	0.309	0.146	0.167	0.551
Horowitz 1951	3.6	3.2	0.554	0.334	0.464	0.158	0.076	0.259
Horowitz 1972	2.1	0.7	0.370	0.174	−0.028	0.537	−0.065	0.283
Iturbi	2.9	5.9	0.784	0.269	0.684	0.178	0.558	0.049
Janis	2.5	4.8	0.755	0.217	0.536	0.193	0.492	0.242
Johannesen	3.3	6.6	0.873	0.428	0.245	0.319	0.569	0.342
Joyce	3.9	5.5	0.725	0.368	0.374	0.177	0.447	0.226
Kahn	3.8	8.2	0.884	0.170	0.464	0.291	0.667	0.352
Karolyi	3.9	4.2	0.779	−0.005	0.475	0.488	0.374	0.470
Katz	3.7	5.3	0.826	0.219	0.548	0.403	0.522	0.164
Kentner	3.6	9.3	0.936	0.324	0.342	0.280	0.624	0.481
Kersenbaum	2.9	5.5	0.862	0.454	0.479	0.209	0.359	0.433
Kilényi	2.9	3.4	0.778	0.409	0.307	0.623	0.241	0.211
Koczalski	3.8	2.9	0.401	0.012	0.776	0.243	0.099	−0.137
Koyama	4.4	4.9	0.801	0.591	0.035	0.439	0.388	0.302
Kyriakou	6.6	7.1	0.887	0.259	0.186	0.551	0.598	0.361
Larrocha	4.0	6.1	0.805	0.459	0.499	0.192	0.399	0.242
Levant	4.7	8.2	0.838	0.364	0.303	0.165	0.682	0.278
Liberace	2.5	4.8	0.658	0.321	0.529	−0.034	0.153	0.539

Pianist	ms.d.	ΔMA	UPC-I	PC-I	PC-II	PC-III	PC-IV	PC-V
Licad	4.2	6.0	0.912	0.471	0.319	0.527	0.418	0.313
Lopes	3.1	5.2	0.859	0.408	0.298	0.292	0.472	0.426
Lortat	3.2	3.1	0.664	0.643	0.249	0.311	0.245	0.025
Lortie	2.8	3.5	0.758	0.350	0.022	0.407	0.277	0.645
Magaloff	2.8	5.6	0.904	0.345	0.313	0.391	0.642	0.284
Magin	3.8	6.5	0.834	0.511	0.056	0.258	0.392	0.608
Malcuzinsky	2.1	2.3	0.433	0.672	0.294	−0.090	−0.037	0.134
Mamikonian	2.5	4.7	0.873	0.491	0.221	0.269	0.530	0.389
Manz	5.4	4.0	0.743	0.404	0.100	0.449	0.167	0.574
Murdoch	3.0	4.6	0.849	0.389	0.581	0.427	0.321	0.226
Niedzielski	3.0	2.9	0.500	0.001	0.047	0.228	0.345	0.482
Novaes	5.2	7.0	0.900	0.347	0.260	0.405	0.668	0.277
Paderewski	2.9	2.7	0.696	0.201	0.321	0.444	0.598	−0.044
Pennario	3.0	7.4	0.866	0.350	0.480	0.053	0.688	0.291
Penneys	3.2	6.9	0.816	0.517	0.282	−0.035	0.483	0.509
Perahia	4.5	2.3	0.727	0.167	0.134	0.678	0.325	0.361
Perlemuter	2.5	4.6	0.649	0.220	0.784	0.255	0.141	0.141
Pokorna	2.7	5.1	0.829	0.505	0.240	0.223	0.290	0.594
Pollini	2.8	6.2	0.915	0.428	0.229	0.347	0.511	0.498
Ranki	3.4	2.7	0.749	0.631	0.114	0.510	0.195	0.238
Renard	3.4	3.4	0.759	0.105	0.395	0.667	0.416	0.159
Richter	2.2	6.6	0.699	0.590	0.620	−0.002	0.306	0.037
Saperton	4.2	5.0	0.718	0.271	0.417	0.227	0.407	0.278
Sasaki	3.6	6.9	0.865	0.581	0.158	0.173	0.397	0.584
Sauer	3.3	6.7	0.854	0.291	0.290	0.194	0.661	0.406
Schein	2.7	6.8	0.867	0.236	0.375	0.159	0.632	0.488
Shebonova	2.6	5.6	0.876	0.447	0.468	0.236	0.312	0.517
Simon	3.3	4.0	0.792	0.216	0.367	0.323	0.392	0.488
Skavronsky	3.8	3.3	0.699	0.416	0.064	0.511	0.131	0.481
Slenczynska '56	5.2	2.8	0.657	0.095	0.414	0.754	0.197	0.107
Slenczynska '75	4.2	1.7	0.592	−0.101	0.355	0.741	0.218	0.209
Slobodyanik	4.2	6.3	0.896	0.391	0.250	0.379	0.408	0.575
Smith	3.3	4.8	0.801	0.594	0.238	0.255	0.147	0.582
Sofronitzky	4.1	5.8	0.853	0.225	0.288	0.390	0.616	0.353
Solomon	3.1	4.8	0.688	0.514	0.138	0.015	0.661	0.081
Székely	4.5	6.9	0.842	0.607	0.123	0.235	0.423	0.442
Timofeyeva	3.2	4.3	0.820	0.716	0.104	0.308	0.301	0.375
Uninsky	3.0	6.7	0.731	0.236	0.661	0.019	0.546	0.144
Varsi	3.5	5.4	0.870	0.308	0.132	0.387	0.615	0.449
Vásáry	2.9	4.9	0.724	0.645	0.234	0.090	0.355	0.243
Vered	3.2	5.0	0.831	0.391	0.255	0.424	0.340	0.462
Virsaladze	3.4	8.2	0.848	0.467	0.297	0.259	0.490	0.347
Volondat	2.3	6.0	0.832	0.502	0.368	0.122	0.612	0.181
Weissenberg	2.9	3.8	0.774	0.498	0.101	0.423	0.373	0.313
Wild	3.9	6.8	0.826	0.482	0.450	0.146	0.450	0.291
Woodward	4.3	2.8	0.711	0.277	−0.040	0.703	0.458	0.174
Woytowicz	2.9	4.8	0.751	0.463	0.548	0.495	0.224	0.005
Yamazaki	2.9	5.6	0.875	0.410	0.319	0.307	0.406	0.509
Yokoyama	3.5	4.0	0.702	0.647	0.013	0.211	0.190	0.486
Zarankin	3.9	6.6	0.871	0.490	0.459	0.275	0.572	0.111
Zarankin 2	3.9	6.7	0.872	0.490	0.461	0.273	0.572	0.112
Zayas	3.9	6.8	0.854	0.487	0.366	0.258	0.302	0.509

[1]Slurs and expression marks are omitted in this computer-generated score, and the second half of bar 5, which was not included in the measurements, has been condensed into a chord to save space. See Repp (1998a, Fig. 1) for an original score of the music.

[2]All timing-dynamics correlations were computed between the IOI durations and the PSLs of the tones *initiating* (and filling) the IOIs. The initial upbeat and the final, very long IOI were not included, so that there were only 35 data points.

[3]Although, in theory, an especially creative artist could go beyond the strategies identified here and choose a completely novel pattern, one that might map onto a "nonsignificant" PC profile in the PCA, such patterns are more likely to be perceived as abnormal and do not seem to occur in the present sample.

[4]Although the statistical support is similar, the argument is weaker for dynamics than for timing because the typical timing pattern is backed up by strong perceptual results (Repp, 1992, 1998b) which have not been dupli-

cated for dynamics so far (Repp, 1995c). The typical timing pattern seems to be "demanded" by the musical structure in a way that the typical (melodic) dynamic pattern is not.

[5]Although the author generally insists on the terminological distinction between *note* (a printed symbol) and *tone* (an acoustic signal), in the context of MIDI applications it is customary to refer to actions on the keyboard as *notes*. MIDI notes have onsets, offsets, and velocities; tones have onsets, poorly defined offsets (because of decay), and peak intensities; printed notes have none of these, only nominal values.

[6]One data point has been omitted from Fig. A3(b) because it seemed anomalous. In position 2-1 of the music [see Fig. A1(b)], a very soft primary alto note was accompanied by a louder bass note; the velocity difference was −13, but the PSL difference was almost 10 dB, which is well above the solid regression line in Fig. A3(b).

Gabrielsson, A. (1987). "Once again: The theme from Mozart's Piano Sonata in A major (K.331)," in *Action and Perception in Rhythm and Music*, edited by A. Gabrielsson (Royal Swedish Academy of Music Publication No. 55, Stockholm), pp. 81–103.

Martin, D. W. (1947). "Decay rates of piano tones," J. Acoust. Soc. Am. 13, 535–541.

Palmer, C. (1996a). "Anatomy of a performance: Sources of musical expression," Music Percept. 13, 433–453.

Palmer, C. (1996b). "On the assignment of structure in music performance," Music Percept. 14, 23–56.

Palmer, C., and Brown, J. C. (1991). "Investigations in the amplitude of sounded piano tones," J. Acoust. Soc. Am. 90, 60–66.

Repp, B. H. (1992). "Probing the cognitive representation of musical time: Structural constraints on the perception of timing perturbations," Cognition 44, 241–281.

Repp, B. H. (1993). "Some empirical observations on sound level properties of recorded piano tones," J. Acoust. Soc. Am. 93, 1136–1144.

Repp, B. H. (1995a). "Acoustics, perception, and production of *legato* articulation on a digital piano," J. Acoust. Soc. Am. 97, 3862–3874.

Repp, B. H. (1995b). "Expressive timing in Schumann's 'Träumerei:' An analysis of performances by graduate student pianists," J. Acoust. Soc. Am. 98, 2413–2427.

Repp, B. H. (1995c). "Detectability of duration and intensity increments in melody tones: A partial connection between music perception and performance," Percept. Psychophys. 57, 1217–1232.

Repp, B. H. (1996a). "The dynamics of expressive piano performance: Schumann's 'Träumerei' revisited," J. Acoust. Soc. Am. 100, 641–650.

Repp, B. H. (1996b). "Patterns of note onset asynchronies in expressive piano performance," J. Acoust. Soc. Am. 100, 3917–3932.

Repp, B. H. (1997a). "The aesthetic quality of a quantitatively average music performance: Two preliminary experiments," Music Percept. 14, 419–444.

Repp, B. H. (1997b). "Expressive timing in a Debussy Prelude: A comparison of student and expert pianists," Musicae Scientiae 1, 257–268.

Repp, B. H. (1997c). "Acoustics, perception, and production of legato articulation on a computer-controlled grand piano," J. Acoust. Soc. Am. 102, 1878–1890.

Repp, B. H. (1998a). "A microcosm of musical expression: I. Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major," J. Acoust. Soc. Am. 104, 1085–1100.

Repp, B. H. (1998b). "Variations on a theme by Chopin: Relations between perception and production of deviations from isochrony in music," J. Exp. Psychol. 24, 791–811.

Repp, B. H. (in press). "Control of expressive and metronomic timing in pianists," J. Motor Behav.

Repp, B. H. (in preparation). "A microcosm of musical expression: III. Aesthetic evaluation of pianists' expression in the initial measures of Chopin's Etude in E major."

Shaffer, L. H. (1981). "Performances of Chopin, Bach, and Bartók: Studies in motor programming," Cogn. Psychol. 13, 326–376.

Todd, N. P. McA. (1992). "The dynamics of dynamics: A model of musical expression," J. Acoust. Soc. Am. 91, 3540–3550.

Todd, N. P. McA. (1994). "The auditory "primal sketch:" A multiscale model of rhythmic grouping," J. New Music Res. 23, 25–70.