1189

Music Perception
Winter 2000, Vol. 18, No. 2, 173–211

© 2000 BY THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
ALL RIGHTS RESERVED.

Pattern Typicality and Dimensional Interactions in Pianists' Imitation of Expressive Timing and Dynamics

BRUNO H. REPP
Haskins Laboratories

Success in an imitation task may be defined in absolute terms (how well the reproduction matches the model) or in relative terms (how much better the reproduction matches the model than the imitator's spontaneous production matches the model). In the present study of expressive music performance, it was hypothesized that absolute imitation success would decrease but relative imitation success would increase as the model pattern diverges more and more from the typical pattern of spontaneous production, within reasonable limits. This hypothesis received support in Experiment 1, which required pianists first to play a musical excerpt spontaneously and then to imitate model performances instantiating different patterns of expressive timing or dynamics, with the other dimension held constant at a typical pattern. The typical expressive pattern for a given musical passage seems to function as a cognitive schema that biases perception, memory, and (re)production of expressive nuances. The results also suggested that imitation of different dynamic patterns affects expressive timing (louder tones were followed by longer interonset intervals), whereas imitation of different timing patterns had little effect on produced dynamics. The latter findings were replicated in Experiment 2, which presented models that differed in both timing and dynamics. Attention to both dimensions simultaneously also reduced imitation accuracy, especially for timing.

Received August 23, 1999, accepted June 20, 2000.

ONE of the deepest and least understood (because hardly investigated) questions in the psychology of music is how artists achieve individuality of expression in their performances of a musical composition. Although this question may be posed in many cultural and stylistic domains, the present discussion will be restricted to the Western "classical" tradition, where a printed score is available. This makes it easier to distinguish between piece-independent and piece-specific aspects of individuality. The

Address correspondence to Bruno H. Repp, Haskins Laboratories, 270 Crown St., New Haven, CT 06511-6695. (e-mail: repp@haskins.yale.edu)

ISSN: 0730-7829. Send requests for permission to reprint to Rights and Permissions, University of California Press, 2000 Center St., Ste. 303, Berkeley, CA 94704-1223.

former reflect technical abilities and performance habits that may be considered personal characteristics of an artist. The latter reflect an artist's engagement with a particular composition and are commonly referred to as interpretation. Interpretation may in turn be divided into two components, one due to cognitive analysis and mental representation of the musical structure (i.e., form), and the other due to feeling and expressive characterization (i.e., content or meaning), although the second component may to some extent be implied by the first.

Psychologists interested in music performance have focused primarily on the cognitive level and thus have regarded expressive nuances as a means of conveying or clarifying musical structure to the listener (see, e.g., Clarke, 1985; Palmer, 1989, 1997; Todd, 1985). Individual differences in interpretation, however, may reflect more often differences in expressive characterization than differences in structural analysis (see, e.g., Gabrielsson & Juslin, 1996; Repp, 1998a; Shaffer, 1995). Structural conceptions are usually limited to a few discrete options, often of very unequal acceptability, whereas expressive characterization draws on a multidimensional continuum of possibilities. In other words, there are many expressive characterizations that are consistent with any given structural conception.

Structural analysis is a skill that is taught explicitly in music schools, although its rudiments can also be acquired implicitly through mere exposure to music. By contrast, expressive characterization is rarely taught systematically. For example, there are hundreds of textbooks on music theory and ear training, but very few on how to play with good expression. One reason for this discrepancy is that the principles underlying expressive characterization are complex and poorly understood; another reason may be that instruction in expressive performance has to be tailored to a much greater degree to students' individual abilities and personalities. Therefore, the necessary skills are usually transmitted from teacher to student through metaphorical discourse and presentation of explicit models for imitation. The abilities to translate a verbal metaphor into movement/sound and to copy a teacher's performance of a short passage are essential to the acquisition of advanced skills on a musical instrument, as any private lesson or master class demonstrates. Even professional musicians often need to imitate each other's expression, for example, when taking turns in playing the same passage in an ensemble or when following a conductor's sung phrasing in rehearsal. However, these abilities have rarely been investigated by psychologists, and they are not included in standard musical ability tests (see Shuter-Dyson, 1999). Some reasons for their absence from ability tests are not difficult to see: Grading a student's success would require either a large panel of judges (in the case of metaphors) or objective measurements (in the case of imitation), both of which are time-consuming and require special expertise. Fortunately, however, such considerations are not an obstacle to empirical investigation of these abilities.

The present study is concerned with imitation of expressive nuances in music performance, leaving metaphor aside for the time being. The music is assumed to be familiar and to present no technical difficulties to the participating musicians.

Imitation requires three basic psychological processes: perception, memory, and action. First, the model performance must be perceived accurately. Auditory discrimination thresholds, perceptual biases, and limits of attention play a role at this level. The perceptual information must then be retained in some form until imitation begins. Even if reproduction starts as soon as the model ends, the beginning of the model has occurred some time ago and must be retrieved from memory. As the actions required for imitation are carried out, the remembered details of the model must be retrieved continually to guide the imitation. Limits to the span and accuracy of memory will play an important role at this stage, as will the extent to which the perceptual information lent itself to recoding into some more abstract and economical representation that can be "unpacked" during reproduction. As to the reproduction itself, it may be biased by cognitive schemata or ingrained action patterns.

The main hypothesis underlying the present research was that imitation of expressive models will be biased by the patterns of expression that the imitator would adopt spontaneously. Previous research has shown that young pianists' spontaneous expressive patterns tend to be close to the typical or average pattern for the music at hand (Repp, 1995b, 1996, 1997, 1998a, 1999a). At least in the case of expressive timing (the pattern of tone interonset intervals, also called rubato), such a typical pattern is evident (albeit often in attenuated form) even when the pianist's intention is to play without expression (Behne & Wetekam, 1993; Drake & Palmer, 1993; Palmer, 1989; Penel & Drake, 1998; Repp, 1999b, 1999d), and the perception of performance timing exhibits biases that mirror the typical expressive timing pattern (Repp, 1992, 1998b, 1998c, 1998d, 1999d). In other words, the typical timing pattern seems to be demanded by the musical structure, as it were. Therefore, it was expected that pianists' perception, memory, and reproduction of different expressive timing patterns would be biased in the direction of the typical timing pattern. An analogous prediction was made for the imitation of expressive dynamics (the pattern of relative tone intensities), although that prediction was more tentative because the perception and production of expressive dynamics have been less investigated and seem less constrained by a typical pattern (Repp, 1995a).

The imitation of expressive dynamics does not seem to have been investigated previously. However, pioneering research on pianists' ability to imitate expressive timing has been conducted by Eric Clarke (1993a; Clarke & Baker-Short, 1987; see also Clarke, 1993b), whose findings in part motivated the present work.

The main purpose of Clarke's studies was to demonstrate that a timing pattern is easier to imitate when it is appropriate to the musical structure than when it is not. Clarke and Baker-Short (1987), in their main experiment, presented each of two short monophonic melodies in three versions: "deadpan" (i.e., with perfectly regular timing), "structured" (i.e., with expressive timing appropriate to the musical structure), and "perverse" (i.e., with structurally inappropriate timing). The structured and perverse versions were synthesized by the investigators on the basis of their musical intuitions. Four pianists heard each version three times and then played three consecutive imitations of it. The main results were that, as predicted, the structured versions were imitated more accurately and with greater consistency (i.e., less variability across the three imitation attempts) than the perverse versions, even though the latter were imitated quite well. The deadpan versions were imitated most accurately and consistently, although there was a tendency to deviate from strict regularity in the direction of the structurally appropriate timing pattern, as has been observed by other authors in deadpan or "mechanical" production tasks (e.g., Behne & Wetekam, 1993; Drake & Palmer, 1993; Penel & Drake, 1998; Repp, 1999b, 1999d).

In a more elaborate follow-up experiment, Clarke (1993a) used two longer melodies, each lasting perhaps 20 s, in addition to the short tunes from the earlier study. Each melody was presented in four versions, one being an expressive performance by a pianist (deemed to be structurally appropriate) and the other three having been generated under computer control by applying structurally inappropriate transformations to the appropriate timing pattern: two translations (i.e., phase shifts relative to the metrical grid) and an inversion. Ten pianists heard each version three times and imitated it once after each presentation. A significant improvement was found from the first to the second imitation attempt but not from the second to the third. Again, structurally appropriate timing was imitated more accurately and more consistently than the transformed versions, with the inverted pattern being the most difficult to imitate. In his discussion, Clarke argued that the accurate imitation of a structurally appropriate timing pattern is due to its being generated from a mental representation of the musical structure, whereas the pianists' moderate success in imitating structurally inappropriate patterns cannot be so explained. After dismissing the implausible possibility that such patterns are retained and regenerated with the help of an anomalous structural representation, Clarke concluded that they must be retained in a more literal fashion: in an auditory memory, by means of verbal tags, or perhaps in a motor memory.

It is very likely that Clarke's structurally appropriate timing patterns were also highly typical for his melodic materials. In other words, they were similar to timing patterns that young musicians tend to produce spontaneously. This raises an interesting question: If the participating pianists had simply played the music in their own preferred way, without paying any attention to the structurally appropriate model, would their productions have matched the model just as closely as their presumed reproductions did? And if so, could imitation really be said to have occurred? Clarke did not record any spontaneous performances, so his conclusion that structurally appropriate timing can be generated from a structural representation of the music may be merely a statement of what happens in typical expressive performance, with or without a model. In other words, it is not clear whether the structurally appropriate model actually had any effect on the pianists' behavior, especially with regard to the pattern of timing.

At this point it is necessary to distinguish between two possible definitions of success in an imitation task. According to an *absolute* definition, the measure of success is the similarity of the reproduction to the model. According to a *relative* definition, the measure of success is the extent to which the similarity of the reproduction to the model exceeds the "baseline" similarity of the spontaneous performance to the model. The absolute definition seems appropriate in situations where the relevant behavior is not yet in the imitator's repertoire or where it typically does not resemble any of the models. The relative definition seems more appropriate when the imitator is predisposed to a pattern of behavior that resembles one or more of the models. If music performance were typically deadpan, an absolute definition of imitation success would be appropriate. However, spontaneous performances of music are typically expressive (as Clarke, 1991, acknowledges). Moreover, there is a specific expressive pattern that is most typical, particularly for timing. Here a relative definition of imitation success seems more appropriate.

The absolute and relative definitions of imitation success may also be contrasted in terms of their focus on different psychological processes underlying imitation. The absolute definition is primarily action-oriented, for it stands to reason that a particular behavioral pattern is the easier to (re)produce the more similar it is to the spontaneously produced, typical pattern. The relative definition is more perception- and memory-oriented: If a model differs only slightly from the spontaneous, typical pattern, these differences may not be perceived at all or may be assimilated to the typical pattern in memory (cf. Bharucha & Pryor, 1986; Krumhansl, 1990). Large differences, by contrast, are easier to perceive and remember. In other words, the model specifies to the imitator not so much what to do (he or she knows that already) as what to do differently from the way he or she normally does it.

These considerations led to the following specific prediction: The less similar a model is to spontaneous performance, the less well it will be imitated in absolute terms but the better it will be imitated in relative terms, within reasonable limits. The qualification is necessary because wildly inappropriate expressive models, such as Clarke's "perverse" versions, may not only be difficult to (re)produce but also difficult to remember because they conflict with the musical structure they are attached to. Experiment 1 deliberately stayed within such reasonable limits by eschewing arbitrarily distorted models and instead taking advantage of the results of recent performance analyses (Repp, 1998a, 1999a) that provided a basis for the construction of several different structurally appropriate expressive patterns for the same music.

The models used will be described fully in the Methods section. They differed from each other in three main ways: First, they differed in their similarity to each imitator's spontaneous performance, the principal hypothesis being that relative success in imitation would be inversely related to that similarity. Second, they varied in typicality. Because the participating pianists' spontaneous performances were expected to have highly typical characteristics, it was expected that the models' typicality would be highly correlated with their similarity to the participants' spontaneous performances. Therefore, it was also predicted that relative success in imitation would increase as model typicality decreased. Third, the models differed in complexity. More complex models were expected to be more difficult to remember and hence more difficult to imitate in either absolute or relative terms. Typicality and complexity will be defined more precisely later. The relative importance of these two factors was of interest.

Since all models were derived from analyses of expert performances, they were all considered structurally appropriate, at least more so than arbitrarily distorted models. It is debatable whether they were all equally appropriate. To the extent that structural appropriateness admits of degrees, however, it may be closely related to typicality because structural appropriateness is surely a hallmark of typical performances. As will be seen, there was actually a "touch of perversity" in the models used, particularly in those for expressive dynamics, due to methodological limitations in the construction of the materials. However, it was not considered necessary that the models satisfy the highest aesthetic standards, as long as they retained some significant resemblance to patterns observed in expert performances.

A second theoretical issue of considerable interest concerned the relation between the (re)production of expressive timing and expressive dynamics. Somewhat surprisingly, earlier performance analyses (Repp, 1996, 1999a) did not reveal any significant interdependences between these two dimensions. Therefore, the (null) hypothesis investigated here was that timing and dynamics would also be imitated independently. Experiment 1 provided a weak test of this hypothesis because it used a design in which only one dimension was varied at a time in the models, so that participants could attend selectively to that dimension. The other dimension was held constant at a typical pattern. Here the prediction was that (re)production of the constant dimension would not be affected by the imitation of different patterns of the other dimension. Experiment 2 subsequently subjected the hypothesis of dimensional independence to a stronger test by varying both expressive dimensions simultaneously in a factorial design.

Experiment 1

METHODS

Musical Materials

The musical excerpt was the opening (bars 1–5) of Frédéric Chopin's Etude in E Major, op. 10, no. 3. A computer-generated score without slurs or expression marks is shown at the top of Figure 1. The second half of bar 5 was condensed into a chord to give the excerpt maximal closure.

Figure 1a shows the most typical expressive timing pattern (or profile) for this music (T0). It represents the average profile of 115 expert performances whose timing was measured from digitized acoustic recordings (Repp, 1998a).[1] The graph represents interonset intervals (IOIs) as a function of metrical (score) position. The initial upbeat IOI, corresponding to an eighth note in the score, is excluded in this graph and all following ones of timing; its average duration was 1122 ms. All other IOIs represent sixteenth-note intervals in the metrical scheme. IOIs initiated by melody notes (among other notes) are shown as filled circles, those initiated only by accompaniment notes as open circles. The melodic line, in the highest voice, is divided into six rhythmic groups (runs of filled circles in the graph), each ending with a long note during which the accompaniment in the other voices continues. It can be seen that the T0 profile includes ritardandi within each of the melodic segments, as well as a lengthening of the final IOI in bar 3 (which is the initial IOI of the longest melodic group) and sometimes of the final IOI of an accompaniment passage immediately preceding a melodic group (the initial IOIs in bars 2, 3, and 5).

Figure 1b shows the most typical dynamic profile (D0). It represents the average dynamic profile of the same 115 expert performances (Repp, 1999a) and includes the initial upbeat and the final chord. The original measurements were peak sound levels that were influenced by all simultaneous tones, but mainly by melody tones, if present, because they were the most intense tones. The peak sound levels were converted into MIDI (key-press) velocities, as explained below. It can be seen that melody notes (filled circles), with the exception of the initial upbeat, were played louder than accompaniment notes (open circles), and that a dynamic peak was reached at or just before the melodic pitch peak in bar 3. An initial crescendo and final diminuendo are also evident. Although the accompaniment during sustained melody notes is typically played both faster and softer than the melodic passages, timing and dynamics show little systematic relationship within the melody (Repp, 1999a).

In order to synthesize realistic performance models instantiating these and other expressive profiles, a good performance of the excerpt by a graduate student pianist was taken as the starting point. The performance had been recorded some time ago in MIDI format on a Roland RD-250s digital piano. The MIDI instructions were converted into text and imported into a spreadsheet program. There the onsets of the "primary notes" were identified,

---

[1] Of course, these performances did not have a final chord in the second half of bar 5 and therefore generally showed less of a ritardando and diminuendo in that bar than is observed in laboratory performances of the isolated excerpt (cf. Repp, 1997).

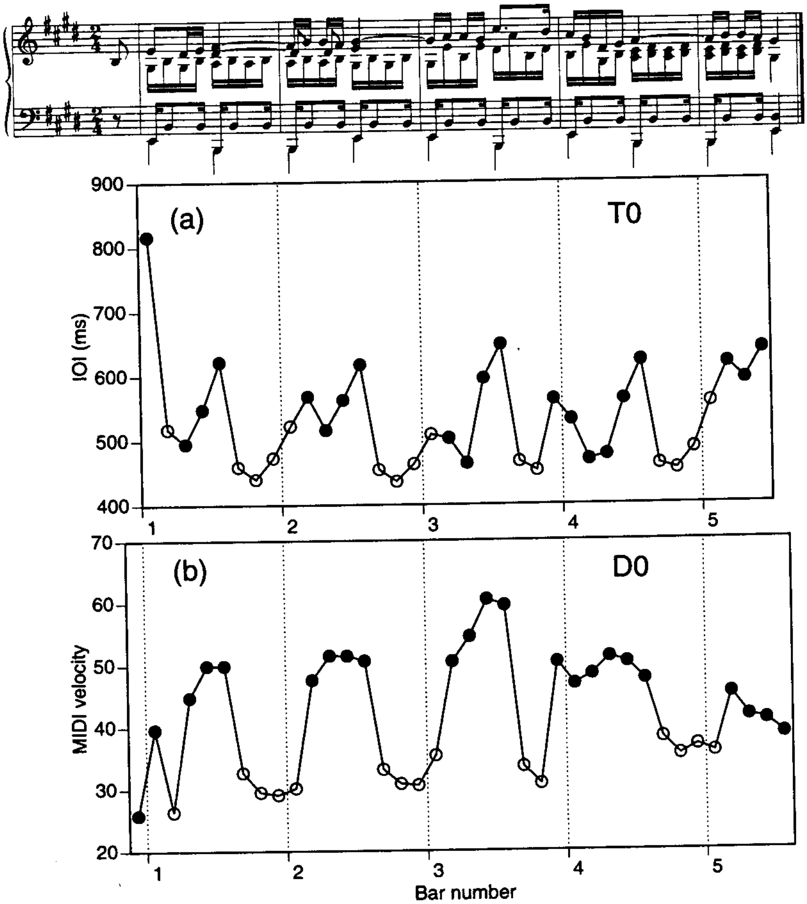

Fig. 1. (top) A computer-generated score of the opening of Etude in E Major, op. 10, no. 3, by Frédéric Chopin. (a) The most typical timing profile (T0). (b) The most typical dynamic profile (D0). IOIs initiated by melody notes are shown as filled circles, those initiated only by accompaniment notes as open circles.

defined as the highest-pitched note in each metrical position (i.e., the melody notes and the highest accompaniment notes during sustained melody notes). The IOIs between successive primary note onsets constituted the timing profile of the performance, and the MIDI (key-press) velocities of the primary notes constituted its dynamic profile. To obtain the models for the present study, different timing and dynamic patterns were substituted for these original profiles. Other aspects of the original performance (note onset asynchronies, note overlaps, relative intensities of secondary notes, pedaling) were adjusted in proportion to the changes in primary-note timing and dynamics.

Specifically, the substitution of a new timing profile was carried out as follows: After calculating the primary-note IOIs of the original performance, the intervals between all MIDI events (including other note onsets, note offsets, pedal depressions, and pedal releases) were expressed as proportions of the IOI in which they were situated. New IOIs were then substituted for the old IOIs, and all interevent intervals were recalculated as proportions of the new IOIs. Finally, the new onset times for all events were obtained by cumulating the new interevent intervals. Thus, the timing of all MIDI events remained relatively invariant within primary-note IOIs.

The substitution of a new dynamic profile was more problematic because the original waveform amplitude envelope measurements represented peak sound levels (in decibels) influenced by all simultaneous tones, not just by the primary tones. However, it would have been too complicated and perhaps impossible to take into account the contribution of secondary tones to the overall amplitude envelope. Therefore, the dynamic measurements were treated as if they reflected the relative intensities of the primary tones only, although that undoubtedly resulted in some slight distortion. (Repp, 1999a, Appendix A, presents data that provide some justification for that shortcut.) The decibel values were first converted to MIDI velocities. Because the models were to be played back on a Yamaha Clavinova CLP-611 digital piano, acoustic measurements were first carried out to determine the relationship between MIDI velocity and peak sound level for midregister single tones ("Piano 1" sound) on that instrument. The relationship was found to be well approximated by the quadratic function, $y = c - 2.91x + .052x^2$, where y is the MIDI velocity, x is the peak sound level in decibels, and c is a constant. The conversion was carried out according to this formula, and the constant was chosen so that the highest MIDI velocity in each model profile was 60. These new MIDI velocities were then substituted for the original MIDI velocities of the primary notes. The MIDI velocities of all other notes were recomputed by first expressing them as a proportion of the original MIDI velocity of the nominally simultaneous primary note and then multiplying that proportion with the new MIDI velocity assigned to that primary note. Thus, the relative intensities of simultaneous notes remained roughly the same as in the original performance. The first model synthesized in that manner, which had both typical timing and typical dynamics (T0/D0), sounded smooth and pleasing to the author when played back on the Yamaha Clavinova, as expected (cf. Repp, 1997).[2]

Six additional models were constructed by keeping the typical dynamic pattern (D0) constant and changing the timing profile only. Three of these new timing profiles (T1, T2, T4) represented the first, second, and fourth Varimax-rotated principal components (PCs) obtained in a principal components analysis (PCA) of the timing profiles of the 115 expert performances (Repp, 1998a).[3] The first, second, and fourth PCs accounted for 31%, 17%, and 11%, respectively, of the variance in the performance data. (T0 represents the first unrotated PC of the PCA, which accounted for 61% of the variance.) Thus, T1 was more typical of expert performance than were T2 or T4, and this was also reflected in their respective correlations with T0 , which may serve as indices of typicality (see Table 1). By definition, the three PC profiles were mutually uncorrelated. They emerged from the PCA in the form of standard scores and were converted into IOIs by multiplying them with the average within-performance standard deviation and adding them to the grand average IOI duration of the 115 performances. Thus they had the same mean (basic tempo) and standard deviation (timing modulation) as T0.

Figure 2 (a, b, c) shows these "pure" PC profiles. The duration of the initial upbeat (not shown), which had not been included in the PCA, was arbitrarily set to 1000 ms in all three models. (Correlations among timing profiles reported in this article never include the initial upbeat IOI.) T1 is characterized by strong ritardandi within all melodic groups, but it lacks the other timing features seen in T0. By contrast, T2 shows a striking accelerando in the melodic group of bar 2 and to a lesser degree also in bars 1 and 5, but not at all in bars 3 and 4. T4 shows pronounced between-group *ritardandi* that exceed the within-group ritardandi, as well as a lengthening of the final IOI in bar 3. In terms of complexity, which may be defined as recodability into structure-based rules, T1 seems the least complex and T2 the most complex pattern. Thus the qualitative pattern of T1 can be characterized as "within-

---

[2] All models can be found and played back at the following internet address: http://www.haskins.yale.edu/Haskins/MISC/REPP/IM.html

[3] The third PC was characterized only by a greatly lengthened initial downbeat IOI. Therefore, it was not of particular interest as a model for imitation.

TABLE 1
Typicality and Complexity Indices for the Imitation Models

Model	Typicality	Complexity	Model	Typicality	Complexity
T0	1.00	.53	D0	1.00 (1.00)	.38 (.40)
T1	.67	.31	D1	.71 (.45)	.36 (.33)
T2	.46	.73	D4	.75 (.50)	.51 (.63)
T4	.36	.58	D5	.71 (.55)	.74 (.83)
T12	.79	.54	D14	.88 (.69)	.29 (.32)
T14	.72	.17	D15	.85 (.67)	.46 (.55)
T24	.57	.61	D45	.88 (.78)	.40 (.35)

NOTE—Indices in parentheses are for melodic dynamics.

group ritardandi," that of T4 as "larger between-group than within-group ritardandi," and that of T2 as "within-group accelerandi in bars 1, 2, and 5, but not in bars 3 and 4." Because these ritardandi and accelerandi are nested in a fairly regular periodic grouping structure, a quantitative index of relative pattern simplicity (i.e., degree of periodicity) is provided by the lag-8 autocorrelation of each pattern, which assesses the similarity of timing patterns from bar to bar. A measure of relative complexity may then be obtained by subtracting this autocorrelation from 1. These indices are shown in Table 1, and they confirm the author's intuitive assessments.

Figure 2 (d, e, f) shows three additional "mixed" timing profiles (T12, T14, T24) obtained by pairwise averaging of the three pure PC profiles. The mixed profiles have the same average tempo as the pure profiles, but they are less strongly modulated, because of the averaging. They are representative of some expert performances (see Repp, 1998a) and are in fact more typical than their constituent pure profiles (see Table 1). They are also mutually intercorrelated ($r = .50$ in each case). T14, which is a mixture of the two less complex pure profiles, exhibits a clear periodicity that makes it considerably less complex than the other two mixed profiles (see Table 1).

Another set of six models had the typical timing profile (T0) and varied in dynamics only. Again, there were three pure and three mixed models, although the former ended up being only semipure. They (D1, D4, D5) were intended to represent the first, fourth, and fifth PCs of the PCA on the dynamic profiles of the 115 expert performances (Repp, 1999a), which had accounted for 17%, 20%, and 14% of the variance, respectively.[4] However, when these models were synthesized and played back, they sounded exaggerated and aesthetically unsatisfactory to the author's ears. There were several possible reasons for this, which need not be belabored here. Although the models could have been improved through local adjustments "by ear and hand," a more principled and less subjective method was preferred. After some exploration, it was decided to mix each of the pure dynamic profiles with the typical dynamic profile (D0) in a ratio of 2:1 (i.e., to compute the average of their decibel values with weights of 2/3 and 1/3, respectively). This had the desired effect of attenuating extremes, although the resulting dynamics still contained occasional jarring accents—the "touch of perversity" referred to earlier. For the purpose of imitation, however, this was not considered a serious problem; if anything, these wayward features were expected to be more noticeable and hence more easily imitated. As a result of the weighted averaging, these semipure models had a somewhat smaller degree of dynamic modulation than D0 and consequently a slightly higher average dynamic level, because the models were equated in terms of maximum rather than average MIDI velocity.

---

[4] The second and third PCs accounted for less variance. The numbering of the PCs follows the output of the statistical program used, which did not reorder the PCs after Varimax rotation (Repp, 1999a).

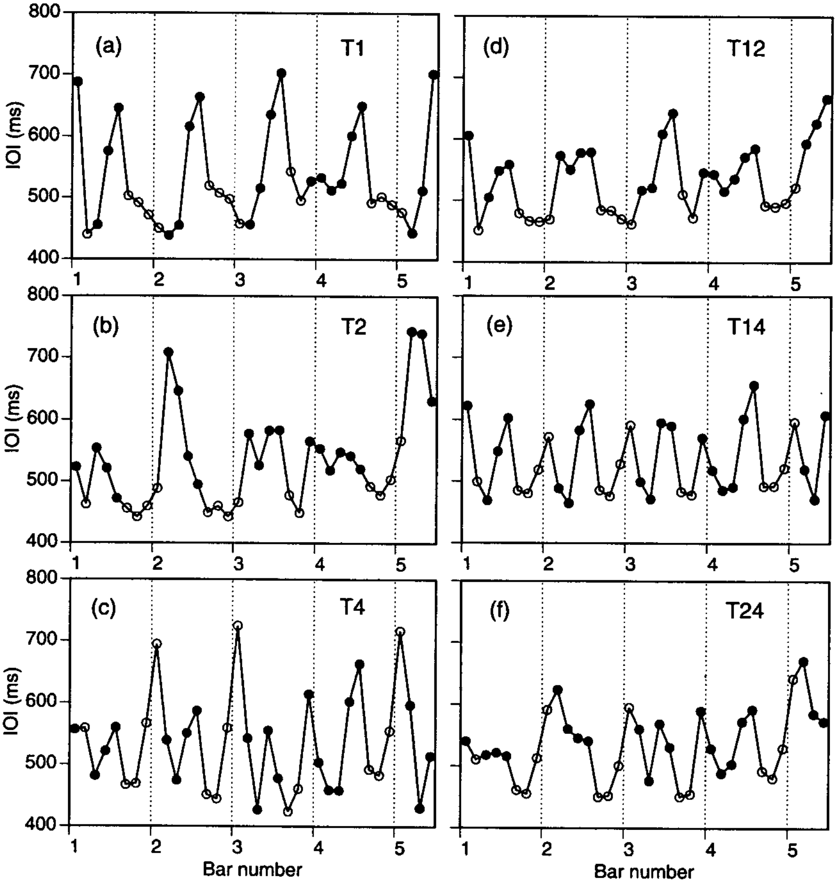

Fig. 2. The less typical timing models used in Experiment 1. (a, b, c) The "pure" models (T1, T2, T4). (d, e, f) The "mixed" models (T12, T14, T24). IOIs initiated by melody notes are shown as filled circles, those initiated only by accompaniment notes as open circles.

The semipure dynamic models are shown in Figure 3 (a, b, c). They are mildly intercorrelated ($r \sim .40$ in each case). Their correlations with D0 are similar (see Table 1); thus the models are of about equal typicality. D1 is characterized by crescendi within melodic groups and absence of an overall arch shape, apart from the very soft initial upbeat. D4, by contrast, is very soft in bar 1 and exhibits sudden reductions in intensity (*subito piano* effects) in the later melodic groups, most notably on the melodic peak in bar 3. D5 shows less dynamic modulation within melodic groups, but instead a dynamic arch building from bar 2 to a peak in bar 4 and then diminishing, with the accompaniment passage in bars 4–5 being as strong as the melody. The complexity index proposed for the timing profiles is less satisfactory here because it only partially captures larger-scale trends extending across melodic groups, but it will do (see Table 1).

Figure 3 (d, e, f) shows the three mixed dynamic models (D14, D15, D45), which were obtained by pairwise averaging of the decibel values of the three semipure models before converting them to MIDI velocities. Owing to the averaging, the mixed models are less modulated than the semipure models, especially within the melody, and their average dynamic level is somewhat higher because of the equalization of maximum velocities. They

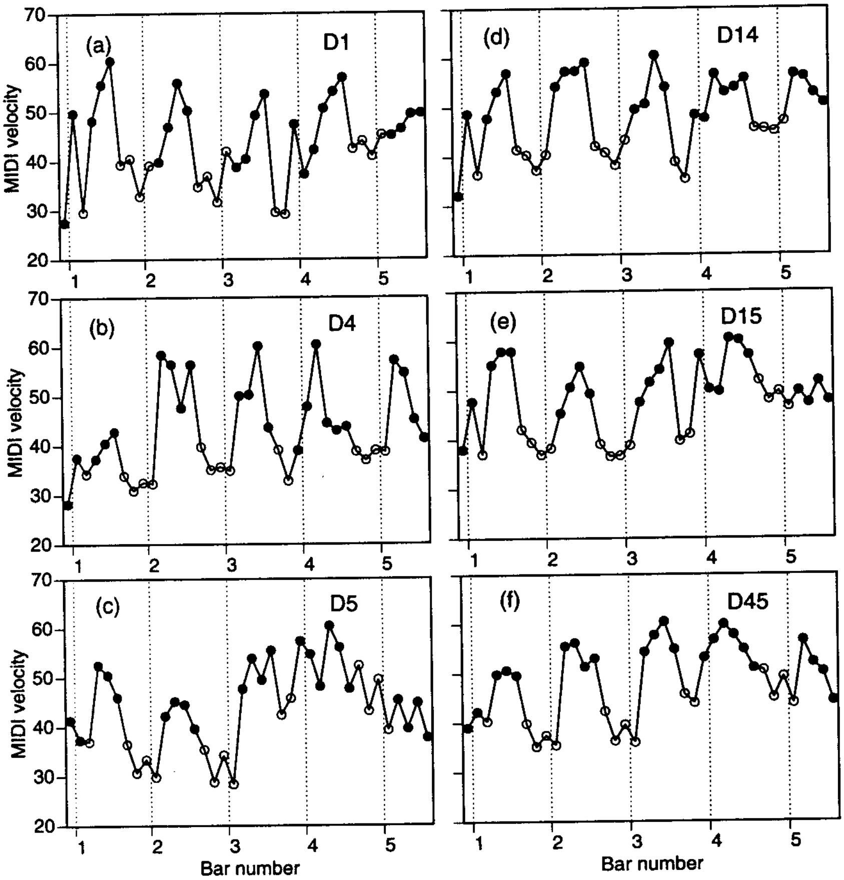

Fig. 3. The less typical dynamic models used in Experiment 1. (a, b, c) The "semi-pure" models (D1, D4, D5). (d, e, f) The "mixed" models (D14, D15, D45). IOIs initiated by melody notes are shown as filled circles, those initiated only by accompaniment notes as open circles.

are more typical than the semipure models (in fact, they are getting fairly close to D0) and do not differ much in relative typicality (see Table 1). Their intercorrelations are .78 to .80, and the differences among them are fairly subtle.

The high intercorrelations of the dynamic models are in part due to the fact that the accompaniment passages are nearly always softer than the melodic passages. Therefore, it may be more appropriate to focus only on the melody notes.[5] In that case, the intercorrelations of the three semipure models range from −.12 to .11 (all n.s.); and those of the three mixed models range from .38 to .61. Thus, the semipure models are very nearly pure (i.e., orthogonal) at the melodic level. The corresponding typicality and complexity indices are shown in parentheses in Table 1.

---

[5] Repp (1999a) also conducted a separate PCA on the melody notes only, with results that were similar to that of the PCA on the complete dynamic profiles.

Participants

Six highly skilled pianists, five women and one man, were paid $25/hour to participate in the experiment. Five of them were graduate students of piano performance at the Yale School of Music (two candidates for M.M.A. degrees, three for M.A. degrees), and one was a senior in Yale college who later went on to piano graduate studies at Juilliard.

Procedure

Because the Chopin excerpt was easy to sight-read, no special preparation was asked for. When a pianist arrived at the laboratory, (s)he was given printed instructions to read and a few minutes to get used to the Clavinova and to practice the excerpt. The score (as shown in Fig. 1, top) was on the music stand throughout the experiment. The sound ("Piano 1") was audible through the built-in loudspeakers of the Clavinova at a comfortable volume. The "touch" of the instrument was set to "hard" (not a real change in key resistance but a reduction of the playing volume) because this reduced an inherent mismatch in volume between MIDI playback and active playing. When the pianist felt ready, (s)he was asked to play the excerpt three times in her/his preferred way while being recorded via a MIDI interface. Recording and playback were controlled by a MAX patch running on a Macintosh Quadra 660AV computer.

Subsequently the MIDI instructions of the first model (which was always T0/D0) were read into the MAX patch by the experimenter. The pianist started the playback of the model her/himself by pressing the highest key on the Clavinova keyboard. After the model had finished, (s)he attempted to reproduce it as faithfully as possible while being recorded. This cycle was repeated twice. The experimenter then saved the data and read in the next model. The next six models either varied in timing only (for three participants) or in dynamics only (for the other three participants). The order of the models was quasi-random and counterbalanced across participants in a Latin square. The pianists were told which dimension (timing or dynamics) varied from model to model and thus should be given primary attention, but they were urged to faithfully imitate the other dimension as well. A short break intervened before the final six models, which varied in the expressive dimension that had been held constant in the previous six models. The whole session took about 1 hr.

RESULTS AND DISCUSSION

Following Clarke's (1993a) example, the expressive profiles of the models and their imitations were compared mainly in terms of their correlations. Correlations capture similarities of pattern, but they do not reflect differences in means (i.e., average tempo or average dynamic level) and standard deviations (i.e., degree of timing modulation or dynamic modulation). These latter differences were considered irrelevant for the purpose of this study. Of course, the IOIs and MIDI velocities of the pianists' spontaneous performances often differed in their mean and/or standard deviation from the models that were presented afterwards, and a closer match to these model parameters was achieved in the imitations. However, the main hypothesis pursued here—that models similar to spontaneous performance should be more difficult to imitate (in the relative sense defined earlier) than models that diverge from spontaneous performance—pertains only to profile similarity as measured by correlations.

Correlational Analysis

Correlations were computed between each model profile and the profiles of each pianist's three spontaneous performances and three imitations of that model, separately for timing, complete dynamics, and melodic dynamics. The average correlations and their standard errors (across pianists) are shown in Table 2. An asterisk indicates that the average correlation of the model with the imitations exceeds the average correlation of the model with the spontaneous performances (the "baseline") by more than the sum of the respective standard errors. (This is only a rough indication of significant differences because it does not take into account that the data were obtained in a repeated-measures design.) In the last column of the table, an index of relative imitation success is reported. It was computed as 100 × $(I_{\max} - S)/(.9 - S)$, where $I_{\max}$ is the largest of the three average correlations obtained in the three imitation attempts (an index of absolute imitation

Table 2

Average Correlations Between Model Profiles and Performance Profiles in Experiment 1, With Standard Errors in Parentheses, and a Relative Index of Imitation Success

Model	Spontaneous	Imitation 1	Imitation 2	Imitation 3	Index
(a) Timing					
T0/D0	.813 (.013)	.819 (.019)	.816 (.012)	.825 (.018)	14%
T1	.784 (.017)	.771 (.029)	.773 (.035)	.781 (.042)	–3%
T2	.257 (.016)	.565 (.037)*	.673 (.034)*	.605 (.021)*	65%
T4	.251 (.025)	.552 (.049)*	.666 (.040)*	.653 (.043)*	64%
T12	.738 (.019)	.798 (.021)*	.778 (.017)*	.799 (.023)	38%
T14	.732 (.016)	.696 (.034)	.732 (.032)	.759 (.029)	16%
T24	.361 (.023)	.554 (.057)*	.608 (.033)*	.613 (.044)*	47%
(b) Dynamics (complete)					
D0/T0	.796 (.029)	.756 (.026)	.817 (.018)	.823 (.017)	26%
D1	.479 (.027)	.603 (.025)*	.706 (.020)*	.649 (.034)*	54%
D4	.556 (.036)	.707 (.030)*	.791 (.023)*	.850 (.013)*	70%
D5	.624 (.032)	.719 (.027)*	.742 (.038)*	.749 (.041)*	45%
D14	.619 (.036)	.756 (.029)*	.810 (.022)*	.783 (.016)*	68%
D15	.654 (.030)	.692 (.038)	.739 (.045)*	.789 (.028)*	55%
D45	.700 (.038)	.719 (.027)	.723 (.038)	.779 (.027)*	40%
(c) Dynamics (melody only)					
D0/T0	.639 (.050)	.466 (.080)	.628 (.041)	.668 (.051)	11%
D1	.146 (.039)	.365 (.094)*	.602 (.036)*	.531 (.068)*	60%
D4	.180 (.060)	.515 (.080)*	.721 (.037)*	.793 (.029)*	85%
D5	.448 (.052)	.603 (.031)*	.679 (.044)*	.692 (.048)*	54%
D14	.220 (.061)	.487 (.075)*	.587 (.086)*	.620 (.041)*	59%
D15	.406 (.050)	.480 (.066)	.570 (.083)*	.630 (.067)*	45%
D45	.441 (.063)	.511 (.048)	.589 (.055)*	.638 (.037)*	43%

*Increase over spontaneous correlation exceeds the sum of the standard errors.

success), S is the average correlation with the spontaneous performances, and .9 is assumed (somewhat arbitrarily) to be the maximal achievable correlation. Thus the relative index expresses how much more similar the imitations were to the models than the spontaneous performances were, as a percentage of the range available for increasing this similarity.

Table 2a reports the results for timing. The first column shows that, as predicted, the pianists' spontaneous timing profiles were highly similar to T0, the typical timing profile. The spontaneous timing profiles were also similar to T1 (the most typical PC profile), T12, and T14, but not to T2, T4, and T24. This was true for all six pianists, as can be inferred from the small standard errors. In absolute terms, T0 was reproduced more accurately than any other model, which is in agreement with Clarke's (1993a) findings. Indeed, absolute success in imitation (I_{max}) increased positively with the baseline correlation across the seven timing models ($r = .96, p < .01$). However, this finding is considered fairly trivial. A more interesting finding is that the correlations between T0 and its imitations did not significantly exceed the baseline. This was also true for models T1 and T14. Thus, although the pianists achieved a significant approximation to all models in the absolute sense (here a correlation of .42 is significant at $p < .05$), there was evidence of significant imitation in the relative sense for only four of the seven models: T2, T4, T12, and T24. This was confirmed ($p < .01$) in separate one-way (spontaneous vs. imitation) repeated-measures analyses of variance (ANOVAs) on the correlations for each model. (These tests were somewhat conservative because they were based on the average model-imitation correlations, not on I_{max}.) Three of the four models showing significant relative imitation success had low baseline correlations. The correlation of the relative imitation indices with the baseline correlations was $-.91$ ($p < .01$), and their correlation with the typicality indices of the models (Table 1) was $-.72$ ($p < .10$). This clearly confirms the prediction that relative success in imitation would be inversely related to the baseline correlation and hence also to the typicality of the model profile.

Although model complexity (Table 1) was negatively correlated with absolute imitation success ($r = -.49$, n.s.), as one might have predicted, it showed a strong *positive* relationship with relative imitation success ($r = .78, p < .05$). This was so because the models with low baseline correlations were also the ones with high complexity indices ($r = -.68$, n.s.). However, the partial correlation of complexity with the relative imitation index was still positive ($r = .54$, n.s.) after removing the effect of the baseline correlation. Clearly, model complexity was not an obstacle to relative success in imitation.

To test whether there was any improvement in the course of the three imitation attempts, a two-way (7 × 3) repeated-measures ANOVA was conducted on all model-imitation correlations, with models and attempts as variables. The main effect of attempts reached significance [$F(2,10) = 4.7$, $p < .04$] and did not interact significantly with models. The average correlations on the three imitation attempts were .679, .721, and .719. Clarke's (1993a) finding that improvement occurred only between the first and second of three attempts was thus replicated.

The T0 model was imitated an additional 18 times in the context of the models with different dynamic profiles. It might be asked whether imitation of this highly typical model improved during these 18 attempts, while attention was focused on dynamics. To address this question, each participant's correlational data were arranged according to the sequential order in which the dynamic models had been presented, and a two-way (6 × 3) repeated-measures ANOVA was conducted on the model-imitation correlations, with the variables of order and attempt. There were no significant effects in this analysis, and the grand average correlation of .803 was actually slightly lower than the baseline correlation and the model-imitation correlations achieved with the T0/D0 model (see Table 2).

Table 2b reports the results for the complete dynamic profiles, and Table 2c those for the dynamic profiles of the melody only. As predicted, the dynamics of the spontaneous performances correlated most highly with D0, the typical dynamic profile. Baseline correlations with D45, D15, and D5 were lower, and those with D14, D4, and D1 were lower still. These differences were more pronounced when only the melody was considered. With regard to imitation, the results were similar in both analyses: In terms of absolute correlations, all models were approximated to some degree, but in terms of the index of relative success, D0 was not significantly imitated. ANOVAs on the correlations for each model also indicated that the average model-imitation correlations for D45 (complete or melodic profile) and D5 (complete profile only) were not significantly higher than the respective baseline correlations. However, this holds only for the average correlations; it can be seen in Table 2 that imitation of these models improved considerably over the three attempts and was significant by the third attempt. Nevertheless, the correlation between the relative imitation indices and the baseline correlations was –.72 ($p < .10$) for the complete profiles and –.87 ($p < .02$) for the melodic profiles. Thus the results for dynamics, like those for timing, confirm the prediction that models similar to spontaneous performance would be difficult to imitate in a relative sense. The correlations between the relative imitation indices and model typicality were likewise negative, –.50 (n.s.) for the complete profiles and –.87 ($p < .02$) for the melodic profiles. The correlations between absolute imitation success and either the baseline or typicality were moderately positive in the case of the complete profiles, but near zero for the melodic profiles.

The correlations between model complexity and imitation success were small for both absolute and relative measures. The partial correlation between complexity and relative imitation success, after removing the effect of the baseline correlation, was –.20 (n.s.) for the complete profiles but an unexpected .85 ($p < .04$) for the melodic profiles. Thus, if anything, complexity facilitated imitation.

An ANOVA on all imitation-model correlations showed a significant increase across the three imitation attempts [$F(2,10) = 29.7$ and 17.8, $p < .001$, for complete and melodic profiles, respectively]. However, as was the case for timing, the improvement occurred mainly from the first to the second attempt. The average correlations for the three attempts were .707, .761, and .774 for complete profiles and .490, .625, and .653 for melodic profiles.

The D0 profile was imitated an additional 18 times in the models that differed from each other in timing. To determine whether any improvement occurred in the course of these 18 trials, each participant's correlational data were arranged according to the sequential order of the timing models, and a two-way (6 × 3) repeated-measures ANOVA was conducted on the model-imitation correlations, with the variables of order and attempt. There were no significant effects. However, the grand average model-imitation correlation (.847 for the complete profiles and .766 for the melodic profiles) exceeded the baseline correlation for D0 as well as the model-imitation correlations for the D0/T0 model (see Table 2). Because of the considerable individual differences among the pianists, however, these differences fell short of statistical significance. Nevertheless, there was evidence that at least some participants imitated the D0 model with some success (in the relative sense) after repeated presentation, even when their attention was focused on timing.

Average Profiles

Although averaging over participants and imitation attempts (as is necessary for concise presentation) smoothes out potentially interesting details, the average profiles do convey some general tendencies. Figure 4 thus compares the model timing profiles (solid line) with the average spontaneous profiles (dotted line) and the average imitation profiles (unconnected data points with error bars). The double standard error bars are equivalent to 95% confidence limits and give a rough indication of where the imitations deviated significantly from the models. In viewing these differences, it should be taken into account that the spontaneous performances tended to be slower in tempo than the models and their imitations, as can be seen most clearly in Figure 4a.

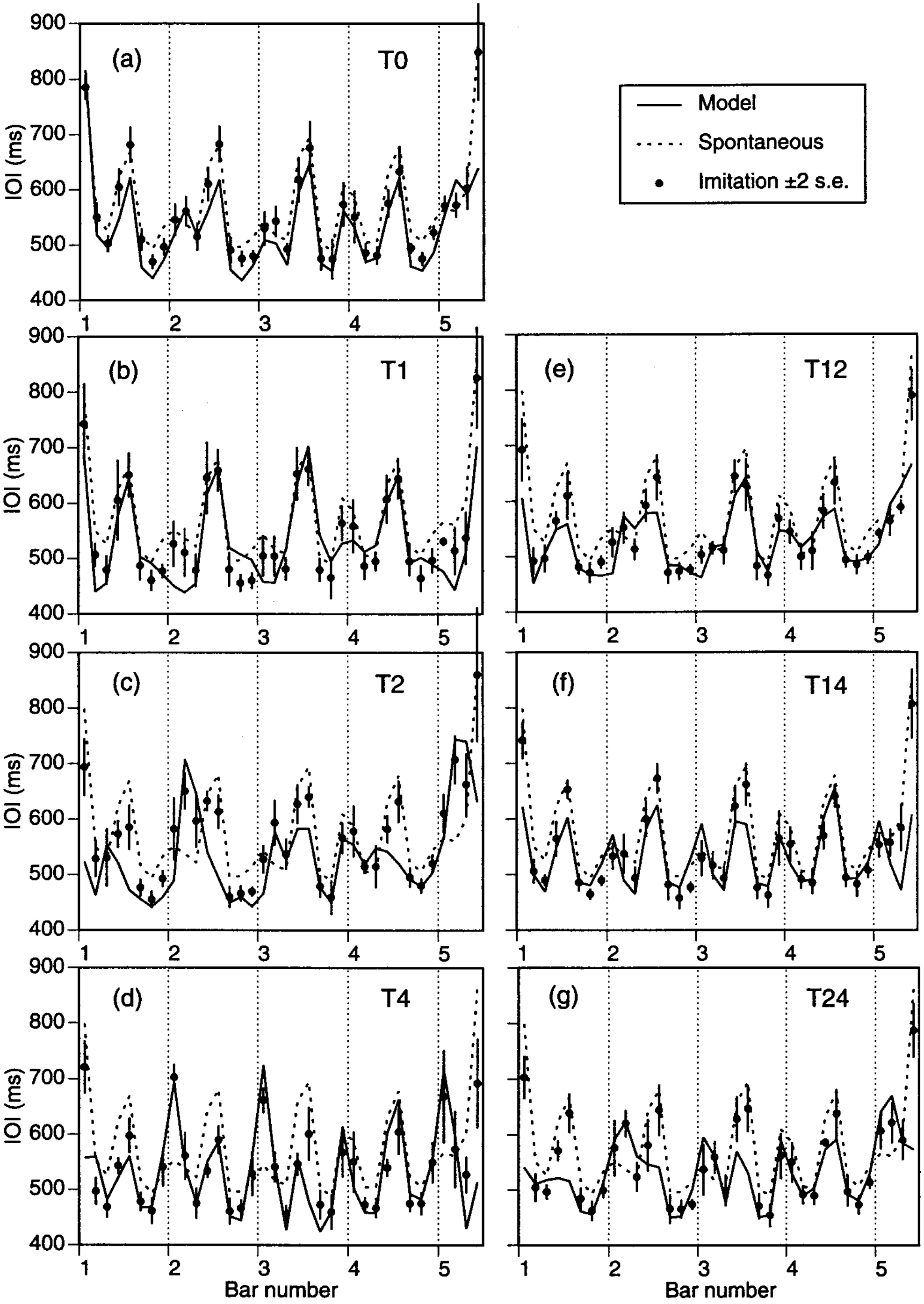

Fig. 4. Model timing profiles, spontaneous timing profiles, and imitation profiles (with double standard errors) in Experiment 1.

It is evident that the largest discrepancies between the models and their imitations occurred in the initial and final IOIs. Apart from the initial upbeat (not shown in Fig. 4),[6] these were the longest IOIs in the spontaneous performances, and in general the pianists played them almost as long in their imitations, even though the models often presented much shorter durations of these IOIs. It must be acknowledged that the final IOI of the models was inappropriately short because the models had been derived from complete performances of the Chopin etude that, unlike the excerpt used in this study, did not end with a chord in bar 5 but continued on with sixteenth-note motion in the accompaniment. However, this does not explain why the final IOI was imitated so poorly. If it had seemed inappropriately short to the pianists, they should have imitated it much more accurately than they did. The results suggest, on the contrary, that the pianists did not notice (or could not remember) the short durations of the initial and final IOIs. Clearly, there were no positive primacy or recency effects in this imitation task.

Model T0 (Fig. 4a) was so similar in pattern to the spontaneous performance that there was little room (though somewhat more at the individual level than is suggested by the average profiles) for improving the pattern match, except in the final IOI, where no improvement occurred. The pianists also tended to overshoot the peaks in the T0 model profile, in accord with their spontaneous performances.

With model T1 (Fig. 4b), there was little room for improved pattern matching at the profile peaks (the ritardandi within melodic groups), but the model differed substantially from the spontaneous timing pattern during the accompaniment passages between melodic groups. Here the pianists did not succeed in suppressing their spontaneous tendency to slow down near the beginnings of melodic groups (beginnings of bars 2, 3, and 5). They apparently did not notice or remember that the model accelerated rather than slowed down in these places. The spontaneous peak straddling the line separating bars 3 and 4 was also maintained, even though it was absent in the model.

Turning next to model T14 (Fig. 4f), which did have between-group ritardandi as well as reduced within-group ritardandi, it can be seen that imitations tended to undershoot the former and overshoot the latter. The short penultimate IOI was completely ignored.

---

[6] The duration of the initial upbeat was 1122 ms in the T0 model and 1000 ms in all other timing models. Its average duration in the spontaneous performances was 1024 ms, but there were very large individual differences (703 to 1527 ms; see also Repp, 1998a). Its average duration was 1229 ms in imitations of the T0/D0 model, 1146 ms in imitations of T0 models with other dynamic profiles, and 1071 ms in imitations of models with different timing profiles. The differences from the model durations were generally nonsignificant because of large variability.

Model T4 (Fig. 4d) had greatly exaggerated between-group ritardandi and greatly reduced within-group ritardandi, both of which were imitated quite well, except at the melodic peak (fifth position in bar 5). The short penultimate IOI was again overshot.

The most salient feature of model T2 (Fig. 4c) was the slow beginning and following accelerando in the melodic group of bar 2, which is also present to some degree in bars 1 and 5. The pianists imitated the slow beginnings in bars 2 and 5 but did not accelerate, continuing instead at a slow tempo in accordance with the peaks in their spontaneous performance. Similarly, they were not able to suppress their within-group ritardandi in bars 1, 3, and 4. Similar problems occurred with model T24 (Fig. 4g) and, to a lesser extent, with model T12 (Fig. 4e).

It appears from these comparisons that the major difficulty in imitation of timing was the suppression of group-final lengthenings when they were reduced or absent in the models. By contrast, it seemed easier to imitate lengthenings that exceeded those in spontaneous performance or did not occur spontaneously at all. Evidently, exaggerations attracted more attention than did attenuations. Such a bias was expected on the basis of very similar asymmetries observed in an earlier perceptual study using the same musical excerpt (Repp, 1998c). The findings are thus consistent with a perceptual explanation of the shortcomings in imitation.

The dynamic model profiles and their average imitations are shown together with the average spontaneous dynamic profiles in Figure 5. It can be seen that the overall dynamic level of the D0 model (Fig. 5a) was close to that of the spontaneous performances, whereas that of the D1, D4, and D5 models (Fig. 5b, c, d) was slightly higher, and that of the D14, D15, and D45 models (Fig. 5e, f, g) was higher still. This difference was due to the way the models had been generated. The pianists evidently adjusted quite well to these overall level changes in their imitations. However, this is a "main effect" that is of less interest here than the "interactions"—the changes in profile shape.

Across all models, with the exception of D5, there was a strong tendency to play the initial upbeat too loud in the imitations, sometimes even louder than in spontaneous performance (models D0 and D45). The final tones of the model, however, were generally imitated accurately.

The major variation in the dynamic profiles was due to the level difference between melody and accompaniment, which was present in all models and often did not deviate much from that in the spontaneous performances. The largest discrepancy between models and spontaneous performances occurred in the second half of bar 4, where the accompaniment was louder in most models. There was considerable variability in the imitation of that feature, and sometimes undershoot occurred (model D15, Fig. 5f). Relative to the preceding melody notes, the pianists generally pro-

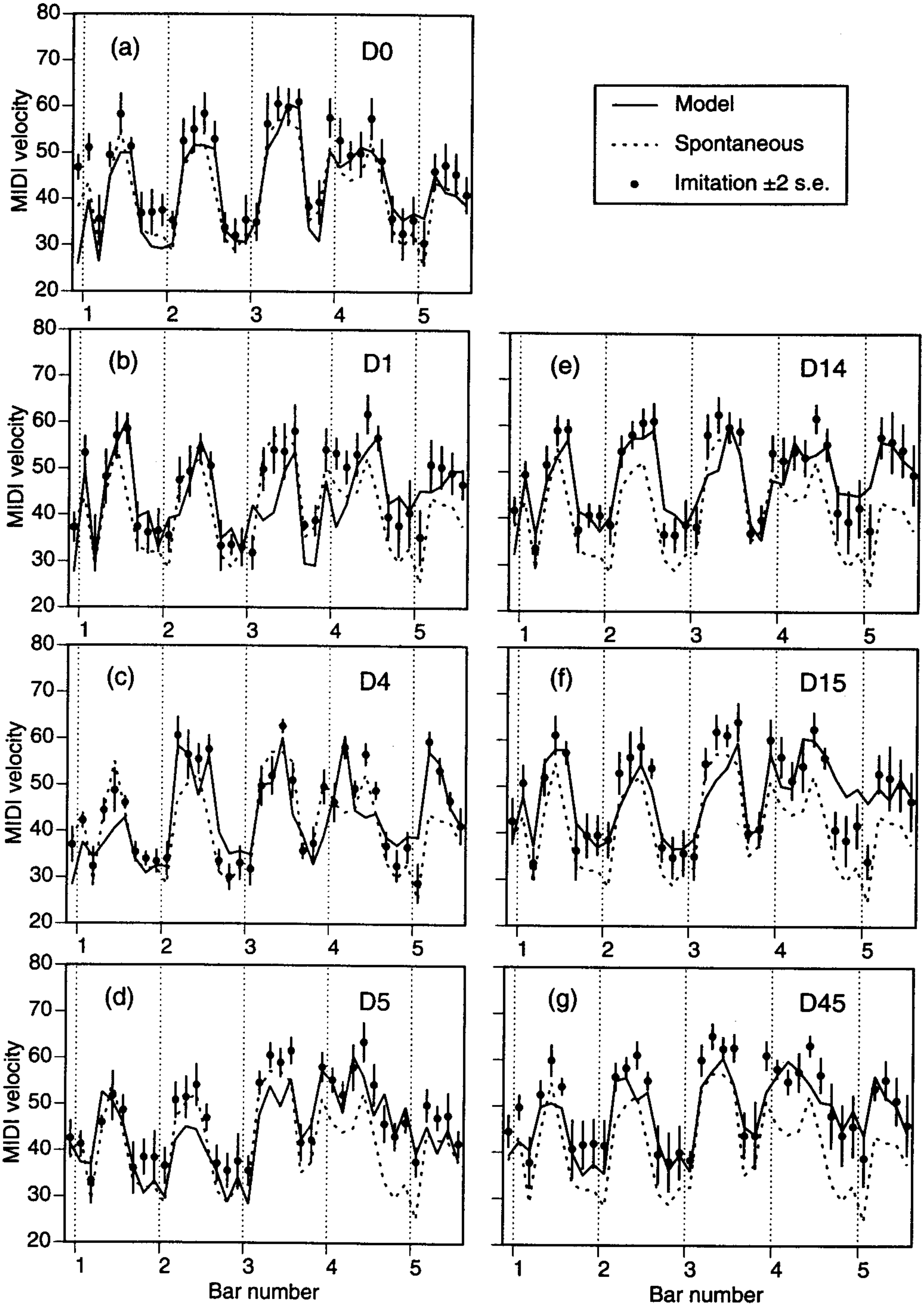

Fig. 5. Model dynamic profiles, spontaneous dynamic profiles, and imitation profiles (with double standard errors) in Experiment 1.

duced a much larger dynamic reduction in the accompaniment than the model exhibited. This was in part due to their inability to suppress the melodic amplitude peak in bar 4, which nearly always overshot the model, most strikingly so in model D4 (Fig. 5c). A similar problem can be seen in bar 1 with most models.

The pianists seemed to be fairly sensitive to the global dynamic contour of the melody, although the imitations often exceeded the model melody in dynamic level. This occurred, for example, in bars 3 and 4 of model D1 and in bars 2 and 3 of models D5, D15, and D45. The detailed dynamic shapes of individual melodic groups were less well imitated, although there was clear evidence of change in response to the models. For example, the *subito piano* effect on the melodic peak (bar 3, position 5) in model D4 was noticed and imitated fairly well. The peak in bar 4, position 2, of the same model is a wayward accent (a "touch of perversity"); it was imitated very accurately, obviously because it attracted attention.

On the whole, these results reveal moderate sensitivity to dynamic differentiation and shape at both gross and detailed levels. In contrast to timing, however, no clear evidence exists for an asymmetry in response to added versus suppressed accents in the models. Both seemed to be imitated to some extent.

Figure 6 gives a summary of absolute imitation accuracy in this experiment, averaged across all models and all participants. The average difference between imitation and model timing profiles is plotted in Figure 6a. This difference profile bears a remarkable similarity to the pianists' average spontaneous timing profile (Fig. 4a, $r = .87$, $p < .001$), which in turn was highly similar to the T0 profile (Fig. 1a), except for a longer final IOI in the spontaneous profile. Thus, despite some success in imitating various aspects of the models, the pianists tended to persist in their preferred, typical timing pattern. In particular, they were reluctant to reduce expressive lengthenings; all large average errors were overshoots of melody IOIs. Figure 6b correspondingly shows the average difference between imitation and model dynamics. Although less regular, this difference profile, too, correlates significantly with the pianists' average spontaneous dynamic profile (Fig. 5a, $r = .72$, $p < .001$), which in turn closely resembles the D0 profile (Fig. 1b). However, the correlation was in part due to a general tendency to exaggerate the difference between melody and accompaniment; the correlation was lower ($r = .49$, $p < .05$) for the melody alone. The exaggeration may be related to the reduction of dynamic contrast in the models, which was due to averaging of dynamic patterns in model construction. In that sense, the exaggeration, too, may represent a tendency to persist in typical patterns of expression.

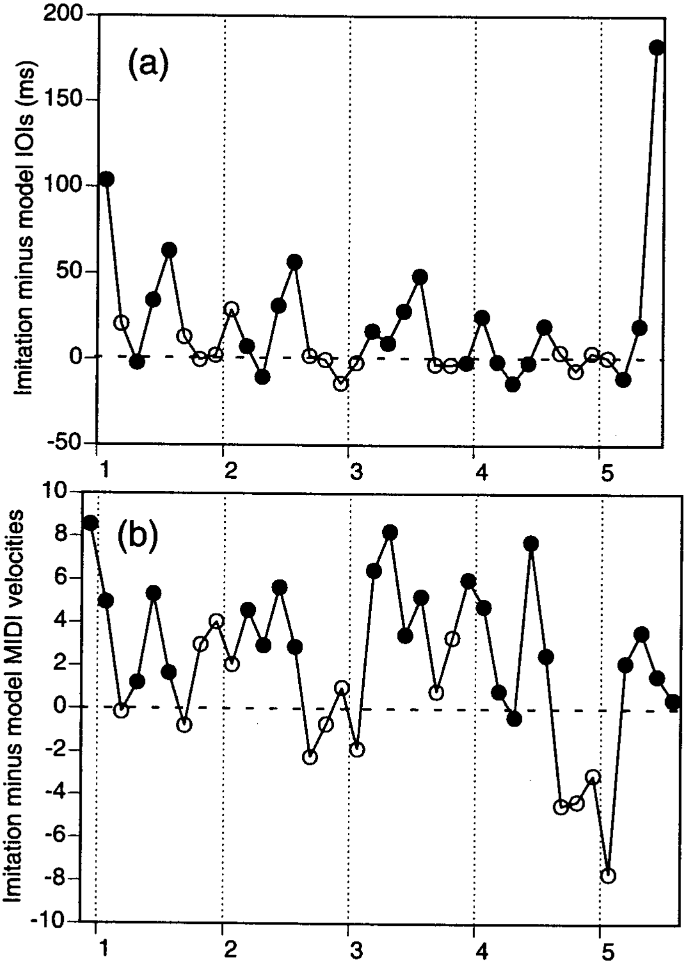

Fig. 6. (a) Average difference between imitation and model IOIs in Experiment 1. (b) Average difference between imitation and model MIDI velocities in Experiment 1. IOIs initiated by melody notes are shown as filled circles, those initiated only by accompaniment notes as open circles.

Interactions between Timing and Dynamics

To test whether variation in dynamics affected the production of timing and vice versa, the following analyses focused on the (semi)pure models only, as they were likely to show the largest effects, if any, on production of the other expressive dimension. Two-way (3 × 3) repeated-measures ANOVAs with model and attempt as variables were first conducted on the correlations between the model profile that was held constant (i.e., either T0 or D0) and the corresponding imitation profiles. An interaction between the two expressive dimensions would here be revealed in a significant main effect of model. The effect of dynamic model imitation (D1, D4, D5) on the typicality (correlation with T0) of the produced timing was nonsignificant [$F(2,10) = 2.1, p < .17$]. The effect of timing model imitation (T1, T2, T4) on the typicality (correlation with D0) of the produced dynamics showed a tendency toward higher correlations when T1 was being imitated, compared to T4 or T5, but that difference likewise fell short of significance [$F(2,10) = 3.6, p < .07$, for the complete dynamic profile; $F(2,10) = 3.8, p < .06$, for the melodic dynamic profile]. Thus, these analyses provided no clear evidence for interactions between timing and dynamics.

However, it could be that the patterns of timing or dynamics changed in some systematic way without that change being reflected in the correlations with T0 and D0, respectively. Therefore, additional three-way (3 x 3 × n) repeated-measures ANOVAs were conducted on the profiles themselves, with the variables of model, attempt, and either IOI ($n = 36$) or MIDI velocity ($n = 38$ or 23). Here the interaction between model and either IOI or MIDI velocity was of interest. That interaction was highly significant for IOI [$F(70,350) = 2.4, p < .0001$], suggesting that the timing profile (imitation of T0) was indeed affected by the imitation of different dynamic profiles; the triple interaction with attempt also reached significance [$F(140,700) = 1.3, p < .02$]. The corresponding interaction for MIDI velocity did not reach significance in the case of the complete profiles [$F(74, 370) = 1.3, p < .09$] but did so for the melodic profiles [$F(44,220) = 1.5, p < .04$].

Figure 7 shows the average timing and dynamic profiles compared in these ANOVAs. Figure 7a suggests that the differences among the timing profiles were primarily due to D4 as compared to D1 and D5. In the first two positions of bars 2 and 5, in position 4 of bar 3, and in position 2 of bar 4, the produced IOIs were longer when D4 was imitated than when D1 or D5 were imitated. The opposite tended to be the case in position 1 of bar 1 and position 5 of bar 3. A comparison of the D4 model profile (Fig. 3b) with the D1 and D5 model profiles (Figs. 3a and 3c) reveals that the former positions generally are associated with peaks in D4, whereas the latter positions are associated with valleys in D4. Thus, increases in the relative intensity of tones seemed to cause a relative lengthening of the following IOIs. This was confirmed by correlating the pairwise differences among the average timing profiles (Fig. 7a) with the pairwise differences among the corresponding dynamic model profiles (Fig. 3a, b, c). The correlations were .61 ($p < .001$) for D1 versus D4, .46 ($p < .01$) for D4 versus D5, and .38 ($p < .05$) for D1 versus D5. Perhaps more appropriately, analogous correlations were computed between the pairwise differences of the timing profiles and the pairwise differences among the average *imitated* dynamic profiles (Fig. 5b, c, d); they were .68 ($p < .001$), .21 (n.s.), and .37 ($p < .05$), respectively. It seems more plausible that timing was influenced

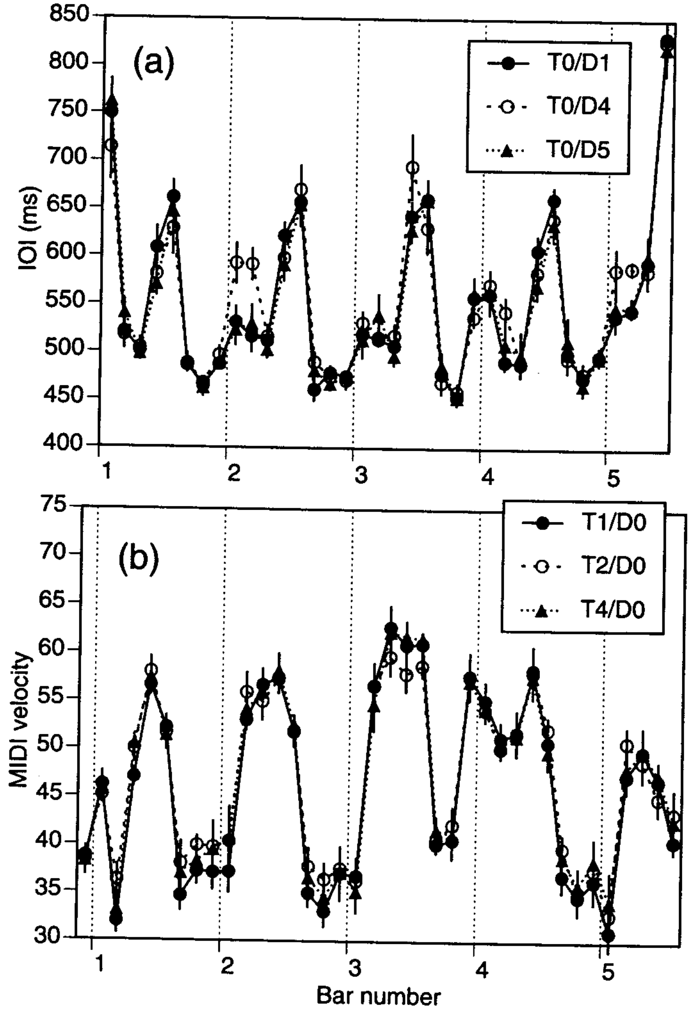

Fig. 7. (a) Imitations of the T0 model when combined with three different dynamic models in Experiment 1. (b) Imitations of the D0 model when combined with three different timing models in Experiment 1. The error bars represent ±1 standard error.

by the produced rather than by the perceived dynamics, though the correlations cannot resolve this issue.

The differences among the dynamic profiles in Figure 7b are smaller and show no obvious relationship to the shapes of the timing models (Fig. 2a, b, c). The correlations between the pairwise differences among the produced dynamic profiles and the pairwise differences among the corresponding timing model profiles (Fig. 2a, b, c) or imitated timing profiles (Fig. 4b, c, d) were all nonsignificant. Thus the effects of different timing patterns on the produced dynamics seemed to be spurious.

Experiment 2

By holding one dimension constant while the other dimension was attended to, and moreover holding it constant at a highly typical pattern, Experiment 1 provided only a weak test of the null hypothesis that there are no interactions between the two expressive dimensions, timing and dynamics. Therefore, a second imitation experiment was conducted in which the models represented orthogonal combinations of different patterns of timing and dynamics, both of which had to be attended to and imitated simultaneously. This design also made it possible to ask whether one of the two dimensions would be imitated more accurately than the other when they are competing for attention, and indeed whether there is any competition for attention at all. Such competition would be reflected in lower absolute and relative indices of imitation success than were obtained in Experiment 1.

METHODS

Musical Materials

Nine new models were constructed by combining the three pure timing patterns (T1, T2, T4) with the three semipure dynamic patterns (D1, D4, D5) in all possible combinations. The method of model construction was the same as in Experiment 1. Model T0D0 (= T0/D0) was also used. The models will be referred to as T0D0, T1D1, etc.

Participants

Six pianists at the same skill level as those of Experiment 1 participated and were paid $25/hour. There were four women and two men; five were candidates for master's degrees and one was expecting an artist's diploma in piano performance from the Yale School of Music. One of them had participated in Experiment 1 almost exactly 1 year ago; it was considered unlikely that she remembered any of the model patterns.

Procedure

The procedure was exactly the same as in Experiment 1. After recording the spontaneous performances, model T0D0 was presented, followed by the 9 new models. The order of these 9 models was different for each pianist and roughly counterbalanced across pianists, such that each model differed from the immediately preceding model in both timing and dynamics. For two pianists, one model had to be repeated at the end of the session due to an earlier failure to record.

RESULTS AND DISCUSSION

Correlational Analysis

The average correlational results are displayed in Table 3 in a format analogous to that of Table 2. The baseline correlations (first column) were quite similar to those obtained in Experiment 1. The average baseline tim-

TABLE 3
Average Correlations Between Model Profiles and Performance Profiles in Experiment 2, With Standard Errors in Parentheses, and a Relative Index of Imitation Success

Model	Spontaneous	Imitation 1	Imitation 2	Imitation 3	Index
(a) Timing					
T0D0	.716 (.051)	.765 (.019)	.767 (.019)	.737 (.031)	28%
T1D1	.698 (.055)	.643 (.036)	.669 (.036)	.640 (.053)	–14%
T1D4	.698 (.055)	.525 (.093)*	.602 (.049)	.552 (.040)*	–48%
T1D5	.698 (.055)	.637 (.046)	.620 (.031)	.587 (.042)*	–30%
T2D1	.275 (.029)	.413 (.083)*	.433 (.038)*	.472 (.043)*	31%
T2D4	.275 (.029)	.427 (.056)*	.528 (.055)*	.530 (.037)*	41%
T2D5	.275 (.029)	.387 (.050)*	.465 (.084)*	.454 (.083)*	31%
T4D1	.278 (.054)	.543 (.019)*	.556 (.029)*	.562 (.027)*	46%
T4D4	.278 (.054)	.412 (.048)*	.489 (.032)*	.434 (.052)*	34%
T4D5	.278 (.054)	.470 (.071)*	.586 (.042)*	.609 (.069)*	53%
(b) Dynamics (complete)					
T0D0	.828 (.019)	.812 (.025)	.860 (.013)	.842 (.011)	44%
T1D1	.510 (.038)	.624 (.017)*	.618 (.029)*	.701 (.011)*	49%
T2D1	.510 (.038)	.651 (.032)*	.680 (.032)*	.659 (.037)*	44%
T4D1	.510 (.038)	.660 (.029)*	.662 (.026)*	.722 (.021)*	54%
T1D4	.540 (.021)	.714 (.026)*	.765 (.027)*	.792 (.018)*	70%
T2D4	.540 (.021)	.661 (.054)*	.728 (.020)*	.780 (.021)*	67%
T4D4	.540 (.021)	.712 (.032)*	.749 (.031)*	.754 (.020)*	59%
T1D5	.656 (.032)	.682 (.025)	.745 (.033)*	.716 (.044)	36%
T2D5	.656 (.032)	.692 (.017)	.708 (.014)*	.742 (.031)*	35%
T4D5	.656 (.032)	.690 (.021)	.723 (.030)*	.748 (.029)*	38%
(c) Dynamics (melody only)					
T0D0	.668 (.053)	.580 (.028)	.722 (.040)	.712 (.024)	23%
T1D1	.156 (.087)	.401 (.044)*	.449 (.048)*	.526 (.044)*	50%
T2D1	.156 (.087)	.441 (.073)*	.518 (.057)*	.520 (.068)*	49%
T4D1	.156 (.087)	.543 (.060)*	.521 (.030)*	.637 (.039)*	65%
T1D4	.098 (.045)	.478 (.070)*	.625 (.057)*	.666 (.048)*	71%
T2D4	.098 (.045)	.419 (.110)*	.583 (.038)*	.683 (.036)*	73%
T4D4	.098 (.045)	.464 (.073)*	.542 (.087)*	.572 (.046)*	59%
T1D5	.484 (.047)	.605 (.041)*	.624 (.033)*	.627 (.052)*	34%
T2D5	.484 (.047)	.436 (.055)	.582 (.028)*	.555 (.080)	24%
T4D5	.484 (.047)	.501 (.046)	.591 (.071)	.599 (.047)*	28%

*Increase over spontaneous correlation exceeds the sum of the standard errors.

ing correlations for T0 and T1 (Table 3a) were somewhat lower and more variable than in Experiment 1, a difference attributed to one pianist who showed much lower correlations than the others because of a fairly mechanical spontaneous performance. That pianist was also the only one who approximated T0 more closely when imitating the T0D0 model, which is in accord with the inverse relationship between the baseline correlation and relative imitation success observed in Experiment 1. The somewhat elevated average model-imitation correlations for the T0D0 model and the relative success index of 28% are due to this one exceptional case; they are not significantly above the baseline. Surprisingly, the T1 model was less well approximated by the imitations than by the spontaneous performances. In separate ANOVAs on the T1 models, however, the differences between the average baseline and model-imitation correlations fell short of significance. The T2 and T4 models, by contrast, showed some significant imitation success, as confirmed in the ANOVAs ($p < .05$ or better). Nevertheless, the imitation indices were a good deal lower than those obtained in Experiment 1, despite similar baseline correlations. This suggests that the presence of different dynamic profiles and/or the simultaneous attention to timing and dynamics interfered with the imitation of timing (see below).

The dynamic model-imitation correlations for the T0D0 model (Tables 3b and 3c) were somewhat higher than the baseline correlation but not significantly so, just as in Experiment 1. The D1 and D4 models, which had lower baseline correlations than D0, especially when only the melody was considered, showed significant imitation in all cases, with the relative success indices being somewhat lower than those obtained in Experiment 1. The D5 model, which had an intermediate baseline correlation, also yielded indices lower than those in Experiment 1, and the ANOVAs based on the average correlations actually fell short of significance, although some individual pianists clearly had more success than others. It appears that the variation in the other expressive dimension and/or the simultaneous attention to dynamics and timing interfered also with the imitation of dynamics.

To confirm these impressions, a 2 × 2 mixed-model ANOVA was conducted on the average model-imitation correlations of Experiments 1 and 2, which seems fair in view of the similar baseline correlations in the two experiments. The ANOVA had the variables of experiment and dimension (timing vs. dynamics). Only the data for (semi)pure models were included from Experiment 1. The correlations for each pianist were first averaged across the three imitation attempts, the three models for each dimension, and the three versions of each model in Experiment 2. The ANOVA revealed a significant main effect of experiment [$F(1,10) = 19.4$, $p < .001$] and also a significant Experiment × Dimension interaction [$F(1,10) = 6.0$, $p < .04$]. Thus, overall absolute imitation success was indeed lower in Experiment 2 than in Experiment 1, and more so for timing than for dynamics.

Interactions Between Timing and Dynamics

To determine whether absolute imitation success for one dimension varied as a function of the other dimension, 3 × 3 repeated-measures ANOVAs with the variables of model (the other dimension) and attempt were conducted on the model-imitation correlations for each of the three timing models and each of the three dynamic models (both complete and melody only), leaving aside the T0D0 model. These analyses also assessed whether there was any improvement across the three imitation attempts. The main effect of model reached significance in only one of the 9 ANOVAs, and then barely [$F(2,10) = 4.2$, $p < .05$]; that was for the T4 model, whose imitation was worse when it was combined with D4 than with D1 or D5 (see Table 3a). The main effect of attempt was significant ($p < .001$) only for model D4, both complete and melodic; in several other cases (T2, D1, D5), the effect merely approached significance ($p < .10$). The interaction of model and attempt reached significance in only one case, for D1 [$F(4,20) = 3.0$, $p < .05$]; as can be seen in Table 3b, imitation did not seem to improve in the presence of the T2 profile.

Even though these results do not suggest much reliable improvement over the three imitation attempts, the results for all models combined (including T0D0) resemble those obtained in Experiment 1: The average model-imitation correlations increased from .522 to .572 but then decreased to .558 for timing, increased from .690 to .724 to .746 for complete dynamics, and from .487 to .576 to .610 for melodic dynamics. Thus, at least for timing and melodic dynamics, most or all of the improvement occurred again between the first and second attempts. Note also that, just as in Experiment 1, the reproduction of melodic D0 dynamics was well below the baseline in the very first imitation trial (Table 3c), perhaps reflecting adjustment to the instrument.

An analysis of correlations is not very sensitive to subtle dimensional interactions, and therefore ANOVAs were also conducted on the actual expressive profiles of the imitations. These were $3 \times 3 \times 3 \times n$ repeated-measures analyses with the variables of primary model (the 3 models whose imitation profiles are being analyzed), secondary model (the 3 models for the other expressive dimension), attempt, and either IOI ($n = 36$) or MIDI velocity ($n = 38$ or 23). The main effects of primary model and IOI or velocity and their interaction were not of interest, as they merely reflected the facts that the imitation profiles were not flat and differed from each other. The effects of interest were those involving the secondary model.

In the analysis of the timing profiles, there was a highly significant interaction of secondary (i.e., dynamic) model and IOI [$F(70, 350) = 3.8$, $p < .0001$] as well as a triple interaction of these two variables with primary (i.e., timing) model [$F(140,700) = 1.7$, $p < .0001$]. This indicates that the shape of the produced timing profiles varied with the dynamic model that was imitated at the same time and that the nature of this variation also depended on which timing profile was being imitated. In separate ANOVAs for each timing model, the interaction between secondary model and IOI was highly significant ($p < .0001$) in each case.

Each panel in Figure 8 shows the average imitation timing profiles for one timing model in the presence of the three dynamic models, with (single)

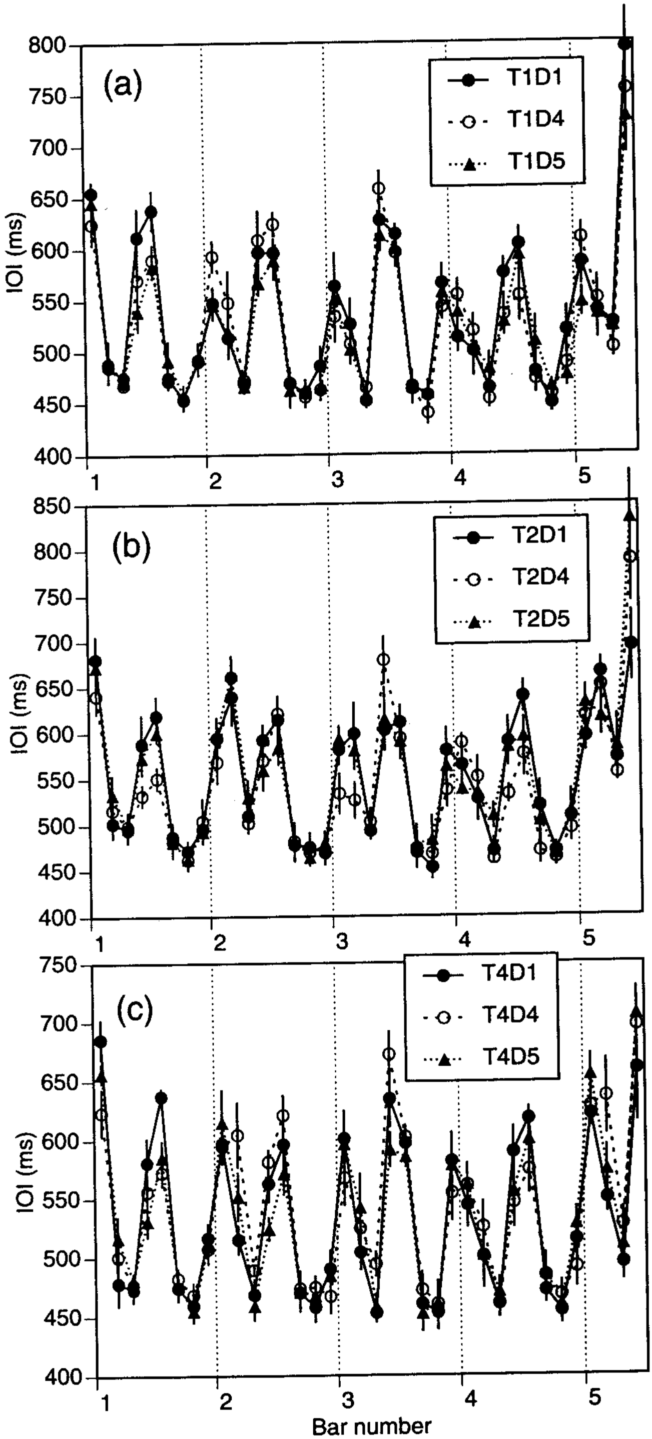

Fig. 8. Imitations of the three timing models when combined with each of three different dynamic models in Experiment 2.

standard error bars. As in Experiment 1, the largest effects on timing appear to be due to the D4 model compared to D1 and D5, and these effects are similar in nature to those observed previously (Fig. 7). It can also be seen that the effects vary in magnitude from panel to panel in Figure 8. Leaving aside these interactions, which are difficult to interpret, the timing profiles were averaged across the three timing models (i.e., across the three panels in Fig. 8) and correlations were computed between the pairwise differences among these average timing profiles and the pairwise differences among either the (complete) dynamic model profiles or the imitated dynamic profiles. These correlations were .70 (D1 vs. D4, $p < .001$), .39 (D1 vs. D5, $p < .05$), and .51 (D4 vs. D5, $p < .01$) for the model dynamic profiles and .63 ($p < .001$), .49 ($p < .01$), and .49 ($p < .01$), respectively, for the imitated dynamic profiles. This replicates the effect of dynamics on timing found in Experiment 1: More intense tones tend to be followed by longer IOIs. The results are also consistent with the assumption that the effects derive from the actually produced dynamics and not from the perceived or intended dynamics.

In contrast to the foregoing analysis of timing, the analogous analyses of the complete dynamic profiles did not yield any significant effects involving secondary (i.e., timing) model, although the triple interaction of secondary model, velocity, and attempt approached significance. In the analysis of the melodic dynamic profiles, that triple interaction was significant [$F(88,440) = 1.5, p< .006$]. In separate analyses on the three dynamic models, however, the triple interaction was significant only for the D5 model [$F(88,440) = 1.4, p < .02$], and the secondary model by velocity interaction also approached significance for that model. Correlations were computed between the pairwise differences among the three average melodic dynamic imitation profiles (after averaging them over the three dynamic model conditions) and the pairwise differences among the three model timing profiles or among the three average imitated timing profiles. The former correlations did not reach significance, but two of the latter did: .44 (T1 vs. T2, $p < .05$), –.11 (T1 vs. T4, n.s.), and .52 (T2 vs. T4, $p < .05$). Thus there was a slight tendency for melodic notes to be played louder when they initiated a longer IOI. Again, this tendency seemed to derive from the IOIs as played, not as perceived. Moreover, the effect is consistent with the effect of dynamics on timing, both showing that longer IOIs go with (preceding) louder tones.

General Discussion

The present study is one of the first investigations of musicians' imitation skills as applied to patterns of expressive musical microstructure. At the most basic level, the results demonstrate that patterns of expressive timing or dynamics, carried by a 20-second musical excerpt, can be reproduced with fair accuracy by skilled pianists, as was shown previously for timing with simpler materials (Clarke, 1993a; Clarke & Baker-Short, 1987). Clearly, performance in this task was far from perfect. If perfect imitation (within the limits of uncontrolled variability) can be achieved at all with such complex patterns, much more practice would be needed than was possible to provide in these experiments. As in Clarke's (1993a) study, there was a significant improvement between the first and second imitation attempts, but little or no improvement between the second and third attempts, especially with regard to timing. This suggests that performance would improve only very slowly with repeated exposure. Moreover, in Experiment 1 there was no significant improvement across 18 additional trials when the dimension imitated was not the focus of primary attention. Given that the unattended dimension was held constant at the most typical pattern, the participants probably just produced their spontaneous profile for that dimension.

Imitation is facilitated when the relevant expressive dimension can be selectively attended to, as a comparison of Experiments 1 and 2 showed. Timing benefited more than dynamics from selective attention, which suggests that dynamic variation intrinsically attracted more attention than timing variation. However, this interaction could well be specific to the materials used, since the dynamic models had some wayward features that were perceptually salient. In general, it is difficult to compare two expressive dimensions with respect to their relative salience. Nevertheless, it is not implausible that dynamic variation is generally more salient than timing variation. For example, Repp (1995a) observed that many individuals have great difficulty detecting small differences in timing, whereas few have such difficulties with small differences in dynamics.

Another asymmetric relationship between timing and dynamics was revealed in both experiments: The production (as distinct from the reproduction) of timing was significantly affected by dynamics, whereas the production of dynamics was barely affected by timing. The effect of dynamics on timing was such that longer IOIs tended to follow louder tones. This is consistent with the observation that the IOIs following accented taps are lengthened in simple finger tapping tasks (Billon & Semjen, 1995; Billon, Semjen, & Stelmach, 1996; Piek, Glencross, Barrett, & Love, 1993; Semjen & Garcia-Colera, 1986). Billon and colleagues have argued that this effect is in part central in origin; in other words, it is not just a consequence of executing a forceful tap more rapidly, but it involves a stretching of the following IOI in the temporal planning of the sequence. The same effect seems to be operating in the context of expressive music performance, where the dynamic levels of all successive notes are more or less different from each other.

It seems plausible that the effect of dynamics on timing was exerted by the dynamic actually produced in imitation of different dynamic models, and not by the perception and memory of the dynamic models themselves. However, it is difficult to distinguish these two possibilities, both of which are consistent with the data. What can be ruled out is the possibility that the model dynamics affected the perceived model timing, which was then imitated as perceived. In that case, the effect should have been reversed because IOIs following louder tones would be expected to be longer and therefore perceived as shorter than they really are. The fact that IOIs following louder tones were lengthened rather than shortened implies that the effect of dynamics on timing production was direct and not mediated by timing perception.

It should be noted that this largely unidirectional interaction between timing and dynamics, although systematic, was fairly small relative to the timing variations that were produced. It also did not significantly affect the relative success in imitating different timing models. Therefore, the results are not in conflict with Repp's (1996, 1999a) earlier findings that suggested that phrase-level expressive timing and dynamics are essentially independent in expert performance, at least in the musical excerpts that were examined. Nor are the results incompatible with Todd's (1992) observation of a negative correlation between IOIs and dynamics at a larger time scale in music, where louder passages are often associated with faster tempi.

The main hypothesis of this study, pursued in Experiment 1, was that relative success in imitation would be inversely related to the baseline correlation, that is to the similarity of the model to spontaneous performance. Because spontaneous performances under laboratory conditions tend to be highly typical (there was one notable exception in Experiment 2 with regard to timing), relative success was also predicted to be inversely related to model typicality. Both predictions were confirmed for both expressive dimensions. The reason for this relationship is believed to be that models are perceived and remembered in relation to the preferred, spontaneous, typical expressive pattern, which is represented internally. If the model differs only slightly from the spontaneous pattern, the differences are difficult to perceive and remember, and they will tend to be assimilated to the spontaneous pattern. If the model differs substantially from the spontaneous pattern, the differences are readily perceived and easier to remember. The spontaneous, typical pattern thus serves as a cognitive schema against which model patterns are evaluated, even in aesthetic judgment (Repp, 1997).

The typical profile also exerted a strong bias on reproduction, especially in the case of timing. In a task requiring the detection of local changes in a typical expressive timing profile for the same Chopin etude excerpt, Repp (1998c) had found that decreases in long (i.e., expressively lengthened) IOIs are more difficult to detect than increases in long IOIs or short IOIs.

Very much the same pattern emerged in imitation of timing, suggesting that the same perceptual factors, perhaps augmented by memory, were at work. In expressive timing there is often a functional asymmetry in that relative lengthenings have expressive importance whereas relative shortenings do not. Therefore, lengthenings tend to be more noticeable than shortenings, as long as the IOIs do not become too short.

It should perhaps be pointed out that, when memory is involved, a cognitive schema can have its effect in at least two ways: by continuously biasing the perception and memory of a model pattern, or by filling in gaps in memory. These two processes are exceedingly difficult to distinguish empirically and may be at work simultaneously.

The present findings are consistent with those of Clarke (1993a; Clarke & Baker-Short, 1987), even though there is a difference of interpretation or perspective. Clarke espoused an absolute definition of imitation success and thus was satisfied to demonstrate that a structurally appropriate (most likely, highly typical) timing pattern is imitated more accurately in an absolute sense than a "perverse" timing pattern. In the present study, too, the most typical pattern was imitated most successfully in an absolute sense. However, this does not seem like a very interesting finding in view of the possibility that presentation of a typical model merely served as a cue to participants to produce their own spontaneous performance. It was argued here that imitation should be considered successful only if the model had a demonstrable influence on the pattern of the imitator's behavior. In the present study, the most typical models did not have such an influence, and it seems likely that they did not either in Clarke's studies. From the perspective of a relative definition of imitation success, which is favored here, Clarke's pianists probably imitated only the perverse models successfully, but not the structurally appropriate model.

Given a negative relationship between the baseline correlation and relative imitation success, it might be asked whether perverse models, such as used by Clarke, would be imitated more successfully than the present, performance-derived, only mildly perverse models. This would require another experiment. It seems likely, however, that the negative relationship demonstrated here would not extend into the region of negative baseline correlations. In that region, perception and memory for model features may not improve further and may even deteriorate because of incompatibility between model features and musical structure. In some sense then, it may still be true that structurally appropriate patterns are easier to imitate than structurally inappropriate ones, even according to a relative definition of imitation success. However, this would have to be demonstrated after first equating the baseline correlations of the models to be compared. Unfortunately, the concept of structural appropriateness remains rather ill-defined. Also, experiments with perverse models are not attractive from an aesthetic viewpoint.

One property of the model patterns that was expected to play a role, if only a secondary one, was their complexity, as quantified by a measure of their degree of periodicity. Surprisingly, complexity not only did not impede imitation but even seemed to facilitate it! This paradoxical effect, which persisted even after the effect of typicality had been partialled out, probably was due to the relatively low complexity of the most typical pattern. Models of high complexity deviated more from the typical pattern, and this contributed positively to perception and memory of the differences. One might expect a more detrimental effect of pattern complexity on imitation when the models are less clearly related to the musical structure or presented outside a musical context.

Even though the present research provides little information about the way in which expressive patterns are retained in memory, it is clear that they are not stored like a list of items that exhibits positive primacy and recency effects. Imitation of timing was conspicuously inaccurate at the very beginning and end of the musical excerpt; imitation of dynamics at the beginning only. The initial and final IOIs are also the ones in which a change in duration is most difficult to detect (Repp, 1998d, 1999c), which is due both to their absolute durations and their adjacency to longer time intervals. This again supports a perceptual explanation of relative imitation success. Clarke's (1993a) very reasonable conjectures about the ways in which expressive information might be retained in memory still await more careful investigation. Obviously, the remembered nuances are somehow pegged to a representation of the musical structure, but exactly how this works remains to be determined.

The present distinction between absolute and relative measures of imitation success was arrived at independently by the author in mulling over Clarke's results. However, it should be mentioned that there is a considerable literature on imitation, primarily in animals and human infants, in which related methodological issues have been discussed at length. For example, Byrne and Russon (1998) distinguish between priming and imitation, where priming is simply the elicitation of behaviors that are already in an organism's repertoire, whereas true imitation gives rise to some novel behavior. This distinction can readily be applied to the present situation: A typical model, which closely resembles a participant's spontaneous performance, may be said to serve as a prime for that performance, whereas an atypical model, which most likely is not in the participant's repertoire, elicits true imitation. Models of different degrees of typicality then may be said to lie on a continuum from pure prime to true stimulus for imitation. Presentation of the music as such may automatically prime the general goal of expressive performance, hence the most typical expressive pattern, resulting in biases toward that pattern in the imitation attempts.

Research on imitation of performance expression may have some relevance to musical performance practice. In his insightful book on jazz performance, Berliner (1994), following up on a quotation from Walter Bishop Jr., distinguishes three stages in jazz artists' development: imitation, assimilation, and innovation. He points out that "musicians who remain at the imitative end of the spectrum enjoy the least prestige" (p. 273). This is equally true in the domain of classical music and is probably the reason why imitation has acquired a bad name. However, this should not detract from the fact that imitation is a necessary first stage in a development that, ideally, should lead to assimilation of the imitated patterns into a rich expressive vocabulary from which new and original patterns and combinations may emerge. Just about everything that Berliner has observed about innovation in jazz performance applies to classical music as well, except that the musical score is sacrosanct, so that all the variation and innovation takes place in the expressive domain.

In today's highly competitive and commercialized world of classical music performance, artists' individuality and originality seems more important than ever. Numerous artists are playing and recording the same standard repertoire, and their success must depend in part on the degree to which their performances are distinguishably different from each other and present new perspectives on the same old music. It is paradoxical, then, that individual differences among performers appear to have shrunk in recent years—a phenomenon that virtually every commentator on the classical music scene has noted. The reasons for this are manifold: the competition circuit, which favors middle-of-the-road interpretations; the obsession with technical perfection, which inhibits spontaneity; the musicological doctrines of faithfulness to the score and "authentic" performance practice, which restrict expressive freedom; the worldwide internationalism, which has eroded local performance traditions; the young age of many performing and recording artists; the increasing substitution of marketing ploys for genuine artistic individuality; and perhaps a general impoverishment of the musical imagination, caused by the urban and technological environment in which most musicians now grow up (Sherman, 1997). By some accounts (e.g., Lebrecht, 1998), classical music is in a crisis, but perhaps increased individuality in performance would provide a partial remedy.

How does individuality come about? To be sure, only few artists are privileged to develop a truly distinctive artistic personality, on purely statistical grounds. Yet this may have been somewhat easier to achieve in the past, due to a greater diversity of cultural and educational backgrounds, and the presence of fewer factors (such as competitions) that discouraged individuality. Today some extra effort may be required to rise above the great mass of technically proficient but musically uninspired performers. Deliberately imitating the less conventional expressive styles and nuances of great artists of the past, not in public performance but as a private training exercise, may hold some promise. It may help develop a larger expressive vocabulary from which the young artist then is free to select whatever he or she finds most appealing. Importantly, it may not be enough just to listen to great performances for inspiration; the active reproduction and embodiment of expressive patterns may be required for assimilation to occur. Imitative exercises of this sort may develop expressive flexibility which then can be exploited according to situational contexts. Conversely, an artist's ability to imitate an expressive model may be an index of the sensitivity and flexibility he or she has already developed. In a way, the musical performer is like an actor who must learn to portray many different characters. Only when that skill has been sufficiently mastered are the many degrees of freedom available that allow an individual personality to emerge in performance.

The models presented in this study, even though they were derived from great artists' performances, were somewhat artificial and flawed. Ideally, original performances should serve as models. Although imitating artificial models may increase flexibility and sensitivity to expressive nuances, imitating original performances by great artists will be a more engaging and meaningful exercise. In that connection, it is interesting to ask what it is that different expressive patterns (such as the ones used as models here) convey. It has already been argued in the Introduction that different expressive patterns often do not convey different structural interpretations. Thus they must represent different characterizations. But what is the character conveyed by each pattern? This question is not easy to answer. To the author's ears, the different timing or dynamic patterns do not convey different moods or emotions; or if they do, there is no vocabulary to describe the very subtle differences.

A metaphor may be helpful here: Expressive performance is like an exhibition of objects in a museum. The musical structure corresponds to the objects, which are arranged in a certain prescribed way. The expression is the illumination brought to bear on the objects, so as to present them in a particular light. One may imagine multiple light sources, various angles, and even colors. Different expressive patterns then may be thought of as different illuminations of the same musical structure; they draw attention to some features and away from others. Lengthened IOIs, increased dynamics, and suddenly decreased dynamics are techniques of guiding the listener's attention through the unfolding structure. Structural interpretation is a special case of illumination in which the musical objects appear to be rearranged. If expressive performance is in part the art of illumination, then imitation of expression may provide a course in lighting design.[7]

---

[7] This research was supported by NIH grant MH-51230. I am grateful to Paul Buechler for extensive assistance and to Eric Clarke, Nigel Nettheim, and two anonymous reviewers for helpful comments on the manuscript.

References

Behne, K.-E., & Wetekam, B. (1993). Musikpsychologische Interpretationsforschung: Individualität und Intention. *Musikpsychologie, 10,* 24–37.

Berliner, P. F. (1994). *Thinking in jazz: The infinite art of improvisation.* Chicago, IL: Chicago University Press.

Bharucha, J. J., & Pryor, J. H. (1986). Disrupting the isochrony underlying rhythm: As asymmetry in discrimination. *Perception & Psychophysics, 40,* 137–141.

Billon, M., & Semjen, A. (1995). The timing effects of accent production in synchronization and continuation tasks performed by musicians and nonmusicians. *Psychological Research, 58,* 206–217.

Billon, M., Semjen, A., & Stelmach, G. E. (1996). The timing effects of accent production in periodic finger-tapping sequences. *Journal of Motor Behavior, 28,* 198–210.

Byrne, R. W., & Russon, A. E. (1998). Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences, 21,* 667–721.

Clarke, E. F. (1985). Structure and expression in rhythmic performance. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition* (pp. 209–236). London: Academic Press.

Clarke, E. F. (1991). Expression and communication in musical performance. In J. Sundberg, L. Nord, & R. Carlson (eds.), *Music, language, speech and brain* (pp. 184–193). London: Macmillan.

Clarke, E. F. (1993a). Imitating and evaluating real and transformed musical performances. *Music Perception, 10,* 317–342.

Clarke, E. F. (1993b). Generativity, mimesis and the human body in music performance. *Contemporary Music Review, 9,* 207–219.

Clarke, E. [F.], & Baker-Short, C. (1987). The imitation of perceived rubato: A preliminary study. *Psychology of Music, 15,* 58–75.

Drake, C., & Palmer, C. (1993). Accent structures in music performance. *Music Perception, 10,* 343–378.

Gabrielsson, A., & Juslin, P. N. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music, 24,* 68–91.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch.* New York: Oxford University Press.

Lebrecht, N. (1998). *Who killed classical music?* Secaucus, NJ: Carol Publishing Group.

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.

Penel, A., & Drake, C. (1998). Sources of timing variations in music performance: A psychological segmentation model. *Psychological Research, 61,* 12–32.

Piek, J. P., Glencross, D. J., Barrett, N. C., & Love, G. L. (1993). The effect of temporal and force changes on the patterning of sequential movements. *Psychological Research, 55,* 116–123.

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. (1995a). Detectability of duration and intensity increments in melody tones: A partial connection between music perception and performance. *Perception & Psychophysics, 57,* 1217–1232.

Repp, B. H. (1995b). Expressive timing in Schumann's "Träumerei": An analysis of performances by graduate student pianists. *Journal of the Acoustical Society of America, 98,* 2413–2427.

Repp, B. H. (1996). The dynamics of expressive piano performance: Schumann's "Träumerei" revisited. *Journal of the Acoustical Society of America, 100,* 641–650.

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

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. *Journal of the Acoustical Society of America, 104,* 1085–1100.

Repp, B. H. (1998b). Obligatory "expectations" of expressive timing induced by perception of musical structure. *Psychological Research, 61,* 33–43.

Repp, B. H. (1998c). The detectability of local deviations from a typical expressive timing pattern. *Music Perception, 15,* 265–290.

Repp, B. H. (1998d). Variations on a theme by Chopin: Relations between perception and production of deviations from isochrony in music. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 791–811.

Repp, B. H. (1999a). A microcosm of musical expression: II. Quantitative analysis of pianists' dynamics in the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America, 105,* 1972-1988.

Repp, B. H. (1999b). Control of expressive and metronomic timing in pianists. *Journal of Motor Behavior, 31,* 145–164.

Repp, B. H. (1999c). Detecting deviations from metronomic timing in music: Effects of perceptual structure on the mental timekeeper. *Perception & Psychophysics, 61,* 529–548.

Repp, B. H. (1999d). Relationships between performance timing, perception of timing perturbations, and perceptual-motor synchronization in two Chopin preludes. *Australian Journal of Psychology, 51,* 188–203.

Semjen, A., & Garcia-Colera, A. (1986). Planning and timing of finger-tapping sequences with a stressed element. *Journal of Motor Behavior, 18,* 287–322.

Shaffer, L. H. (1995). Musical performance as interpretation. *Psychology of Music, 23,* 17–38.

Sherman, R. (1997). *Piano pieces.* New York: North Point Press.

Shuter-Dyson, R. (1999). Musical ability. In D. Deutsch (Ed.), *The psychology of music,* (2nd ed., pp. 627–651). San Diego, CA: Academic Press.

Todd, N. P. [McA.] (1985). A model of expressive timing in tonal music. *Music Perception, 3,* 33–58.

Todd, N. P. McA. (1992). The dynamics of dynamics: A model of musical expression. *Journal of the Acoustical Society of America, 91,* 3540–3550.