ELSEVIER

Cognition 74 (2000) 1–32

COGNITION

www.elsevier.com/locate/cognit

Skill acquisition in music performance: relations between planning and temporal control

Carolyn Drake[a,*], Caroline Palmer[b]

[a]CNRS and University of Paris V, Paris, France
[b]Ohio State University, Colombus, OH, USA

Received 18 December 1998; received in revised form 28 May 1999; accepted 14 September 1999

Abstract

We investigated the acquisition of music performance skills in novice and expert pianists. Temporal disruptions in novice performances coincided with constraints in planning capacities. Child and adult pianists ranging in age (9–26 years), training (3–15 years) and sight-reading ability learned to perform a novel musical piece in eleven practice trials. Computer-detected pitch and timing errors revealed: (1) gradual improvements in performance tempo and pitch accuracy with skill level and practice, generally fitting a power function; (2) a relative-timing/pitch accuracy trade-off and high incidence of simultaneous pitch/time errors; (3) improvements in relative timing (temporal continuity, underlying beat, metrical structure) with skill and practice; and (4) increased anticipatory behavior and a greater range of planning with skill and practice. A strong positive relationship between the mastery of temporal constraints and planning abilities within performance suggests that these two cognitive indicators are closely related and may arise from segmentation processes during performance. Examination of sequence timing may explicate planning abilities that underlie many complex skills. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Music performance; Still acquisition; Planning; Temporal control

1. Introduction

Studies of expert and novice performance in many domains reveal both quantitative differences (such as accuracy and speed) and qualitative differences (such as

* Corresponding author. Present address Laboratoire de Psychologie Expérimentale, UMR 8581 Universite Rene Descartes, Centre Universitaire de Boulogne, 71 avenue Edouard Vaillant, 92774 Boulogne-Billancourt, Cedex, France.
E-mail address: drake@idf.ext.jussieu.fr (C. Drake)

0010-0277/00/$ - see front matter
PII: S0010-0277(99)00061-X

different error types that arise, and sensitivity to different structural aspects). Skilled performance in tasks such as typing, chess playing, and sports is usually faster, more accurate, and reflects more highly structured representations of the task than performance by less accomplished individuals (Ericsson & Smith, 1991; Ericsson & Staszewski, 1989). To a certain extent, these differences between experts and novices can be generalized to a complex task such as music performance. For example, skilled pianists can play faster, make fewer errors, and perform more consistently (McKenzie, Nelson-Schultz & Wills, 1983), and plan ahead (anticipate) more than beginners (Palmer & Drake, 1997). However, the structure of musical sequences imposes additional constraints on performance, namely temporal constraints, which have been severely underestimated in previous descriptions of skilled performance. Temporal constraints in music performance include the need to produce events in a fluent or continuous manner, maintain production rate or tempo, and preserve certain durational relationships between events. In this paper we examine the cognitive bases of music performance with an emphasis on the role of temporal constraints, and their relation to how performers plan, or prepare sequences for production.

We approached the study of music performance in terms of skill acquisition. One reason for this choice stems from the fact that the temporal structure of music tends to be completely mastered in expert performance, providing few instances of temporal deficits or breakdowns. As this study demonstrates, one major component of novices' music performance involves the learning of temporal structure, and performances by beginner musicians provide naturalistic conditions in which to observe temporal breakdowns. We report a study of how musicians of various skill levels learn to play a novel piece of music, focussing in particular on their ability to control both 'what' they play and 'when' they play it. Music performance offers an excellent domain for study of the cognitive processes responsible for both the production of serial order (which event to produce next) and event timing (when), because the temporal precision seen in music performance and the conceptual structures of rhythm and meter indicate that timing is integral to performance goals (for a review see Palmer, 1997).

The acquisition of performance skills can be studied within several paradigms. One approach is to compare performances of musicians at varying levels of expertise (Drake, Dowling & Palmer, 1991; Ericsson, Krampe & Tesch-Romer, 1993). A second approach is to study the progressive acquisition of skills during the learning of a novel piece (Gruson, 1988). We combined these approaches in this study by examining how novice and expert pianists of varying skill levels learn to play a novel (unfamiliar) piece of music. We traced the process of skill acquisition in two ways: first, we tracked improvements in both quantitative and qualitative characteristics of performance by comparing the frequency and type of errors for five groups of pianists who vary in age, training, and sight-reading ability. Second, we examined the time-course of learning by studying improvements observed while each musician learns to play a novel piece of music over many repetitions. The combination of these two approaches provides an indication of which changes are due to general skill level (long-term learning) and which are due to increased familiarity through specific practice of a novel musical piece (short-term learning). Although music performance usually entails much longer and less structured forms of practice, this early (first) stage of practice allows examination of differences in early learning stages that commonly reflect increased variability in other domains (Ivry, 1996; Newell & Rosenbloom, 1981).

2. Four issues of skill acquisition in music performance

2.1. Traditional measures of skill: fluency = speed and accuracy

Previous research has emphasized improvements in speed and accuracy with practice, both across and within skill levels (Chi, Glaser & Farr, 1988; Gruson, 1988; McKenzie et al., 1983), in tasks such as reading inverted text (Kolers, 1975) and cigar manufacturing (Crossman, 1959). The observed learning curves are usually best fit by a power function, which show a rapid improvement at the beginning of practice, with a gradual slowing in improvement as amount of practice increases (Anderson, 1982; Crossman, 1959; Logan, 1988; MacKay, 1982; Newell & Rosenbloom, 1981; Schneider, 1977). Many theoretical interpretations have been suggested to explain the power function, including increases with learning in the size of, and relationships among, memory units or chunks (see Rosenbaum (1991) for a review).

Skilled performance is thus often depicted in terms of the combination of speed and accuracy, or fluency, that results in a smoothness of production in which the correct actions are performed rapidly and without temporal interruption (MacKay, 1982; Shaffer, 1980). However, studies that use fluency as an indicator of skill have focussed primarily on tasks that require rapid, accurate performance, such as typing, tapping, and problem solving. The only temporal goal of these tasks is to perform as fast as possible while making as few errors as possible. A speed/accuracy trade-off is usually observed under these conditions: faster performances tend to contain more errors (MacKay, 1971; Newell & Rosenbloom, 1981).

We examined whether speed and accuracy increase in music performance in a manner similar to the power law function. We traced improvements in terms of performance tempo (a measure of speed) and error rates (a measure of accuracy) at different skill levels (reflecting long-term practice) over multiple repetitions of the same piece (reflecting short-term practice). We also tested for a speed/accuracy trade-off.

2.2. Additional characteristic of skill: fluency = relative timing and pitch accuracy

Complex tasks such as music performance and speech impose additional task constraints beyond fluency, namely, temporal constraints. Overall speed or performance tempo is less important than producing events at the correct point in time. Musicians can create small expressive timing variations, but their productions must reflect the event duration categories specified in a musical score in Western tonal music, which are usually simple integer ratio relationships. Performance is considered deficient if it does not respect these duration categories. Therefore, it is not surprising that a speed/accuracy trade-off has not been observed in expert adult pianists' performances of memorized sequences under speeded conditions (Palmer & van de Sande, 1993) or under normal (unspeeded) conditions (Palmer & van de Sande, 1995): error rates did not increase in faster performances.

The concept of fluency needs to be modified in the case of music performance and other complex serial tasks with high temporal constraints, such as speech. We enlarge the concept of fluency in music performance to include the relative timing of events. A skilled music performance is fluent, in the sense that correct actions are performed at the correct moment in time relative to surrounding events. An additional characteristic of expertise in music performance therefore may be the ability to master temporal constraints by performing events at the correct time. Because speed is replaced by relative timing as a task goal in music performance, we predict a relative-timing/pitch accuracy trade-off, with performers able to produce either the relative timing at the expense of pitch accuracy or vice versa, instead of a speed-accuracy trade-off.

2.3. Mastering the temporal structure of music: temporal continuity, underlying beat and metrical structure

Western tonal music has a hierarchical temporal structure, in which event durations combine to form basic rhythmic patterns which in turn combine to form higher-order rhythmic patterns (Cooper & Meyer, 1960; Lerdahl & Jackendoff, 1983; Yeston, 1976). Three specific temporal constraints on how music is performed are examined here: temporal continuity, underlying beat and hierarchical metrical structure.

The first constraint, temporal continuity, reflects a requirement to produce events in a continuous fashion without hesitations or temporal interruptions. Skilled performers strive to respect the continuous nature of music, for example, concert pianists can adjust the performance of rapidly executed events to compensate for an error in the timing of earlier events (Shaffer, 1980). However, performances by novice musicians demonstrate frequent interruptions, despite music teachers' explicit instructions to continue playing. For example, novices frequently stop to correct errors, whereas advanced musicians tend not to correct their errors (Palmer & Drake, 1997).

The second temporal constraint, maintaining an underlying beat, reflects the need to produce events at time intervals that maintain recognizable (but not necessarily exact) integer ratio relationships. Western tonal music is composed of event durations that are related by integer ratios: event durations are usually one, two, three, or four times as long as the shortest event duration (occasionally more). Respecting these integer ratios is thought to be accomplished by the perceptual extraction of an underlying beat (Dowling & Harwood, 1986; Drake, 1998; Parncutt, 1993). Sensitivity to durational relationships functions early in life: children over 5 years old are better able to reproduce rhythms constructed around simple integer ratios (1:1 and 1:2) than rhythms with more complex integer ratio relationships (Drake & Gérard, 1989). One limitation in generalizing to naturalistic music performance from these simple tasks, however, is the simplicity of the sequences, which usually contained only five to seven events.

The third temporal constraint, metrical structure, reflects the fact that Western tonal music contains temporal regularities defined over longer time-spans (spanning many beats). Meter refers to a pattern of periodic accent placement that arises from alternating strong and weak beats. People reproduce event sequences that are consistent with a metrical framework (those that are constructed around integer ratios) more correctly than those that violate a metrical framework (Essens, 1986, 1995; Essens & Povel, 1985; Povel, 1981), and the same pattern of event durations (such as short-short-long) is performed with different relative timing when placed in different metrical contexts (Clarke, 1985). Music performances often contain expressive nuances that reflect the metrical structure (Drake & Palmer, 1993; Sloboda, 1983; Shaffer, Clarke & Todd, 1985). Events aligned with metrical strong beats are often lengthened (Gruson, 1988) and played louder (Sloboda, 1983). Errors in child pianist's well-learned performances indicate sensitivity to meter: more corrections occur between metrical measures than at other metrical positions (Palmer & Drake, 1997). Some research suggests that young children are not as sensitive to metrical structure as adults (Bamberger, 1980; Drake, 1993a,b; 1997; Drake & Gérard, 1989; Smith, 1983); however, few have compared childrens' and adults' sensitivity in the domain of music performance.

We addressed the effects of both general skill and specific practice on performers' ability to respect these temporal constraints of music. Maintaining continuous temporal performance, respecting the underlying beat, and sensitivity to metrical structure may improve with both general skill and specific practice.

2.4. Preparation of movement sequences: the ability to plan and the range of planning

Many theories of sequence production assume that performers construct a mental plan prior to performance that specifies the content and ordering of sequence events (Keele & Summers, 1976). The number of events that can be prepared simultaneously is limited by memory constraints, and long sequences are thought to be partitioned into shorter subsequences (or segments) during the construction of mental plans (van Galen & Wing, 1984. Evidence from timing and errors in music performance also suggests that musical sequences are partitioned during planning into shorter segments (Palmer & van de Sande, 1995).

Planning abilities are often reflected in anticipatory or future-oriented behavior. For example, production errors in speech and music performance often reveal events that are anticipated (produced ahead of time). A comparison of beginner and intermediate children's piano performances indicated increased anticipatory behavior: children with 2–3 years more musical training showed more anticipatory relative to perseveratory errors (Palmer & Drake, 1997). Specific practice also leads to more anticipatory behavior. Speakers producing tongue-twisters showed an increased proportion of anticipatory errors with practice, and the anticipatory proportion of errors was predicted by the overall error rate (Dell, Burger & Svec, 1997). Dell et al. (1997) propose a model in which practice affects the strength of connections in mental plans to both present and future events but not to past events. Thus, planning capacities improve with practice as a function of increased memory associations among sequence elements. Whether these predictions of specific practice apply to other sequence production domains such as music performance is unknown.

A related question that arises is the range or scope of planning during sequence production. Many errors, such as substitutions, involve the production of a sequence element in a different serial position from that intended by the performer. These errors indicate an interaction between non-adjacent elements that must be simultaneously accessible during planning of the sequence (Garrett, 1980). We assume here that the range of planning, or distance in number of sequence elements between the interacting items (an error and its assumed source), indicates which elements are simultaneously accessible at any given time. Constraints on the range of planning are evidenced in speech errors, whose interacting elements tend to span structural units (Garrett, 1980). Evidence from music performance errors also indicates structural constraints on the range of planning; phrase structure and serial distance limit the range over which sequence elements interact (Palmer & van de Sande, 1993, 1995), similar to error patterns found in speech (Garcia-Albea et al., 1989). Furthermore, the range of planning, as reflected in production errors, increases from beginner to intermediate levels of music performance (Palmer & Drake, 1997). Whether general skill level or specific practice affect the range of planning in terms of greater serial distances, larger structural units, or other factors, is unknown.

We address the effects of both general skill level and specific practice on planning capacities underlying music performance. Both general skill level and specific practice should lead to mental plans that are more future-oriented (as measured in anticipatory errors), and incorporate events from larger ranges of the sequence (as measured by interacting sequence events in errors).

3. Experimental study of musical skill acquisition

We traced the progress of five groups of pianists varying in age, amount of musical training, and sight-reading ability as they learn a novel musical piece. Each pianist performed a piece of music appropriate to his/her skill level in eleven practice trials. We examined the performance tempo, as well as the number and type of errors (pitch, duration, corrections, and pauses), by comparing each performance with intended events notated in the musical score. Based on the previous literature, we make the following predictions.

First, improvements in both error rates (an indication of accuracy) and performance tempo (an indication of speed) should be evidenced over skill levels and practice trials, probably following a power function. We do not however expect a speed/accuracy trade-off as is usually observed in motor tasks, because temporal constraints other than speed (i.e. relative timing) influence the timing of music performance. Second, we do expect a relative-timing/pitch accuracy trade-off in music performance. Beginner pianists may only be able to perform correct pitches at the expense of timing, whereas more skilled pianists may be able to incorporate both pitch and timing dimensions. We compared the relative frequency of pitch and timing errors as a function of skill level and specific practice. Third, skilled performers may respect the temporal structure (relative timing of events) more than beginners. The ability to maintain temporal continuity, respect the underlying beat, and sensitivity to the hierarchical metrical structure may also improve with general skill level and specific practice. Finally, we expected that both general skill and specific practice should lead to mental plans that are more future-oriented (as measured in anticipatory proportions of errors) and incorporate events from larger ranges of the sequence (as measured by interacting sequence events).

3.1. Method

3.1.1. Participants

Five groups of 12 pianists (total of 60) were recruited through local music teachers' associations in the Columbus, Ohio music community: 48 pianists were children and 12 pianists were adults. Children were between 7 and 16 years of age, and adults were between 20 and 30 years of age. Pianists were assigned to one of five groups based on their age and sight-reading ability, as described in the Procedure (Section 3.14) below.

3.1.2. Stimulus materials

Three short Western tonal musical pieces which varied in difficulty were composed for the experiment by a musician to ensure that they would be unfamiliar. The first four measures of each piece are shown in Fig. 1. All three pieces were composed in a Western classical style, with simple meter, melody, and accompaniment. Piece 1, the easiest piece, was a single-voiced melody of eight measures, in

Fig. 1. First four measures of the three musical pieces used in this experiment.

two four-measure phrases. Pieces 2 and 3 were composed of a melody and two accompanying voices: piece 2 had a chordal accompaniment and was composed of two four-measure phrases, and piece 3 was based on a French folk-tune, composed in an ABA format, with each section lasting 16 measures. The pieces are referred to hereafter as p1, p2 and p3.

3.1.3. Apparatus

Pianists performed the pieces on a computer-monitored upright Yamaha Disklavier acoustic upright piano. Optical sensors and solenoids in the piano allowed precise recording and playback without affecting the touch or sound of the acoustic instrument. The timing resolution was 2 ms for note events, with precision (measured by the standard deviation of onset-to-onset durations during recording) within 0.8% for durations in the range of the performances. Errors were identified by computer comparisons of recorded pitch and duration values with the information given in the musical scores.

3.1.4. Procedure and design

The pianists (and their parents) were first interviewed about the pianists' musical background. They performed a familiar musical piece of their own choosing, to become acquainted with the piano. Then they performed one of the unfamiliar musical pieces. All musicians except the adults and the most advanced children began with the easiest piece (piece 1). If they were able to perform the piece with no errors after two trials, they were given the next most difficult piece to try. This procedure was repeated until the musicians settled on a piece of appropriate difficulty. Once the appropriate piece to be learned was established, the pianists performed it five times. Then they performed a short unrelated task, introduced to prevent boredom (an issue for the youngest pianists). Then the pianists were asked to perform the piece once more and that performance was played back by computer for them to hear. Instructions were then given to emphasize either: (1) the meter, (2) the phrase structure marked on the score, or (3) what the music made them think of. These instructions had no significant effects on results presented here and thus are dropped from further analyses. Each pianist then performed the piece an additional five times, with these instructions. In all, each pianist performed one of the pieces 11 times. At the end of the experiment, the children chose a small gift from a collection prominently displayed in the laboratory, and the adults (or parents) received $8.

3.1.5. Skill classification

Each pianist was first classified into one of three sight-reading levels depending on piece difficulty (p1 = easiest, p2 = intermediate, p3 = hardest). Child pianists were further classified on the basis of their age ('young' = 7–11 years old, 'old' = 12–16 years old). This assignment resulted in five groups of 12 pianists each, shown in Table 1: p1-young, p2-young, p2-old, p3-old, and p3-adults. A design that crossed age and sight-reading (skill) level completely was not possible because the easiest piece was too easy for the older pianists to perform with any errors, and the hardest piece was too difficult for the younger pianists (see Table 1). This design does

Table 1
The five groups included in this study

Age	Difficulty of musical piece		
	Easiest (piece 1)	Intermediate (piece 2)	Hardest (piece 3)
Young (7–11 years)	p1-young	p2-young	
Old (12–16 years)		p2-old	p3-old
Adults (>18 years)			p3-adults

however allow the study of improvements during learning due to age, training, and sight-reading ability.

Mean age and years of training (private piano instruction) for each group are shown in Table 2. The 'old' groups (p2-old, p3-old) were significantly older than the 'young' groups (p1-young, p2-young: $F(1,55) = 74.6$, $P < 0.01$) and the adults were significantly older than the old groups: ($F(1,55) = 487.3$, $P < 0.01$). The two 'old' groups had significantly more years of training than the two 'young' groups ($F(1,55) = 31.5$, $P < 0.01$), and fewer years of training than the adults ($F(1,55) = 109.3$, $P < 0.01$). There were no significant differences in mean age or training between the two 'young' groups (p1-young and p2-young) or between the two 'old' groups (p2-old and p3-old).

3.1.6. Error classification

Errors were identified by computer and coded according to a system similar to those used with speech errors (Dell, 1986; Garrett, 1975), adapted previously for the music performance domain (Palmer & Drake, 1997; Palmer & van de Sande, 1993, 1995). Five types of errors were identified in the performances: pitch errors, duration errors, pitch/duration combinations, corrections, and pauses. Examples of the error types are shown in Fig. 2. The experimenters verified the computer coding of the errors based on two musician listeners' notated transcriptions of all of the performances. The musicians' coding was adopted in the rare cases of disagreement with the computer output (usually concerning the position of corrections), which occurred most often in the single-voice piece (p1-young performances) for which

Table 2
Mean, standard deviations, and range of age and training (in years) for the five groups

Group	Age			Training		
	Mean	SD	Range	Mean	SD	Range
P1-young	9.1	1.2	7–11	3.5	1.5	1.5–5.5
P2-young	10.3	1.1	8–11	4.7	1	1.5–7
P2-old	13.1	1.1	12–15	6.6	2.1	4–10
P3-old	14	1.2	12–16	8.4	1.6	5–11
P3-adults	25.7	2.6	20–30	15.3	3.1	10–20

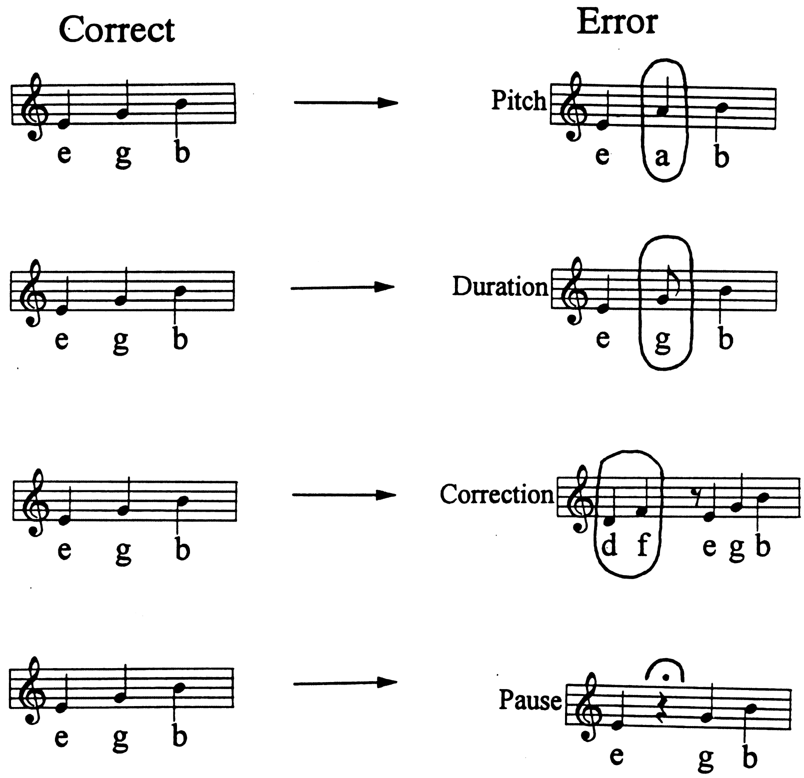

Fig. 2. Examples of error coding system.

error position codings could be ambiguous due to the lack of accompanying voices. Because pause and duration errors are defined here for the first time, the error definitions are described in detail below.

Pitch errors were defined as events whose pitch component differed from the notated pitch information. Pitch errors were coded on the same dimensions as previously (see Palmer & van de Sande, 1993, 1995 for details): type (additions, deletions, substitutions, or shifts), source (contextual, non-contextual, or combination), size (chord, note, or note-chord combination), and movement direction (anticipatory, perseveratory, or exchange). Contextual errors are those whose pitch matches the notated (intended) pitch of a position adjacent to the error. In addition, some errors have a directional component; they can be anticipatory (pitches intended for future positions) or perseveratory (intended for past positions). All errors were coded in the most parsimonious way (to minimize the total number of distinct errors). Following this strategy, multiple pitch errors whose onsets occurred within a small temporal window (75 ms for adults and 100 ms for children) were coded as chord error size; single-note errors were coded as note error size.

Duration errors were defined as events whose performed interonset duration differed from the duration notated in the musical score. (In this paper, duration refers to the interonset interval defined by two successive events.) The observed duration of each event was compared with its predicted duration (based on the duration of the nearest preceding correct event), in order to incorporate local tempo changes. An event duration was considered in error if the observed duration differed from the predicted duration by more than 50% (thereby entering the next duration category in musical nomenclature). The observed duration of the error was then classified into the nearest notational duration category (eighth-notes, dotted eighth-notes, quarter notes, etc.) based on the ratio of observed/predicted durations.

In addition, each duration error was coded for beat disruption: duration errors were considered beat-disruptive if the ratio of the observed (error) to the predicted event duration was not an integer multiple of the beat duration (the beat duration equaled one quarter-note in each musical piece, as defined by the denominator of the time signature). Thus, a quarter-note event performed in error as a dotted quarter-note would be coded as beat-disruptive, whereas a quarter-note event performed in error as a half-note would be coded as non-beat-disruptive. In the case of two adjacent duration errors that occurred within the same beat period (quarter-note), the beat disruption for both errors was coded on the basis of their summed observed duration relative to the (summed) predicted duration.

Pitch-duration errors were defined as events for which both the pitch and duration components were in error. The pitch and duration components of pitch/duration errors were each coded in the same manner as individual pitch errors and duration errors.

Pauses were defined as interonset intervals greater than three times the predicted duration, resulting in values often greater than one second (similar to classifications of hesitation pauses in speech; Fodor, Bever & Garrett, 1974). They may represent silence or a held tone. Predicted durations were estimated from the nearest preceding (correct) event duration, as for duration errors. Pauses were coded as occurring between the serial locations of the event onsets preceding and following the pause.

Corrections were defined as a series of (error) additions of more than one event, followed by a pause, and then a (correct) restart, following previous studies of music performance errors (Palmer & Drake, 1997). Corrections were coded as occurring between the event locations preceding and following the correction.

Learning errors. Some errors in sight-reading tasks may arise from perceptual (reading) or input causes, as well as from production sources (cf. Lehmann & Ericsson, 1993). To separate these sources, we defined learning errors as those pitch or duration errors that occurred consistently in the same event location in four or more of the first five performances. Pitch learning errors formed less than 10% of the total pitch errors and were excluded from the following analyses. In contrast, duration learning errors formed 63% of the total duration errors, 99% of which occurred in performances of piece 1 by the least skilled group (p1-young). Subsequent analyses of these duration errors indicated that some of the least skilled pianists in the P1-young group misinterpreted the time signature as 6/8 instead of the notated 4/4 signature, and performed the event durations to match the incorrect time signature. Thus, the intention rather than the execution was incorrect in these errors. All duration and pitch learning errors were excluded from further analyses.

Error rates. Statistical analyses were conducted on error rates (computed as number of errors relative to number of error opportunities) for each pianist's performances, to adjust for the different lengths of the musical pieces. There was a total of 6089 errors in the 60 (pianist) $\times$ 11 (trial) performances (p1-young = 1685, p2-young = 1277, p2-old = 983, p3-old = 1669, p3-adult = 475). Error rates were computed from a combination of pitch, duration, correction, and pause errors, relative to chance estimates for each error type. Pitch error rates were defined as: [# single note pitch errors ÷ number of single pitches in piece, summing across chords and single-notes] + [# chord errors ÷ number of events in piece][1]. Duration error rates were defined as: # duration errors ÷ (number of events in piece − 1), because there is no event onset to determine the (interonset) duration of the last event. Pitch/duration errors (those containing both a pitch and duration component in error) were divided equally (assigned 0.5) among pitch and duration error rates. Pause error rates were defined as: [# pauses ÷ total events in piece − 1), and correction error rates were defined as: [# corrections ÷ total events in piece − 1]. The mean error rate for each performance was computed from the average of these four numbers. Analyses across all error types were conducted on error rates, and analyses that compare specific error types were conducted on percentage of total errors.

4. Results

4.1. Traditional measures of performance: accuracy and speed

We first examined two hallmarks of skilled performance: increased accuracy and speed, as measured by error rates and performance tempo.

4.1.1. Accuracy

As shown in Fig. 3a, error rates decreased over the five groups and eleven practice trials. An ANOVA on error rates by group (five), practice trials (11), and error type (four) revealed a significant effect of group ($F(4, 55) = 20.7$, $P < 0.01$), with lower error rates for the more skilled groups, and of practice trial, ($F(10, 550) = 19.9$, $P < 0.01$), with error rates decreasing over practice. A significant interaction between practice trial and group ($F(40, 550) = 2.7$, $P < 0.01$) also indicated that the rate at which a new piece improved in accuracy across the eleven trials varied

[1] Pitch error rates can be computed for single-note errors and chord errors (those involving multiple simultaneous note events) by several methods: (1) combine chord and single-note errors and compare total with chance estimate of number of event locations in sequence; (2) treat chords as multiple single-note errors, combine them with number of single-note errors and compare total with chance estimate of total number single pitches in sequence; or (3) compare number of single-note errors with chance estimate of total number of single pitches and number of chord errors with chance estimate of number of event locations; then add these two ratios (each error is represented only once in the two ratios). The first method is biased toward lower error rates when chord errors are uncommon, the second method is biased toward higher error rates when chord errors are uncommon, and the third method typically results in an intermediate value. Given no a priori reason to expect differences in chord versus single-note errors, we use the third method, which represents a balance of the two potential biases.

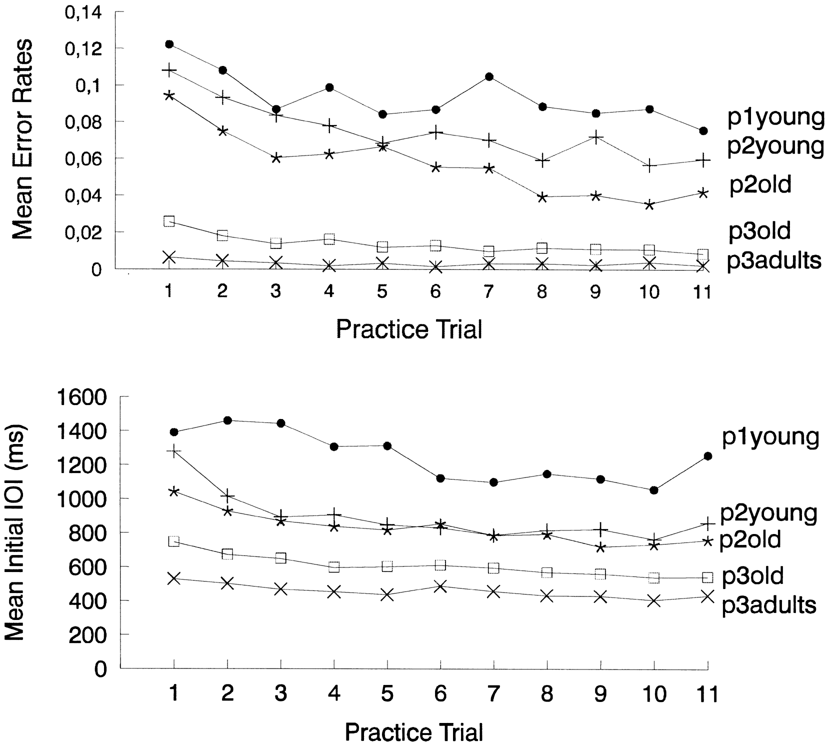

Fig. 3. Improvements over eleven practice trials for the five groups for mean error rates (a) and mean initial IOI (b).

across the groups. To analyze this interaction further, separate ANOVAs were conducted for each group on trials 1–5 and trials 7–11. These comparisons indicated significant improvement over both the first five and last five trials for p1-young (Dunn–Bonferroni, $P < 0.05$), only over the first five trials for groups p2-young, p2-old, and p3-old ($P < 0.05$), and borderline significance over the first five trials for p3-adults ($P = 0.055$). Greatest improvements during practice were found in the least skilled groups.

4.1.2. Tempo

We also examined changes in tempo with learning. The initial tempo at which the pianists performed the pieces was estimated by the mean quarter-note beat duration (IOI) of events in the first error-free measure for each performance (the quarter-note represents one beat, as indicated by the notated time signature[2]). This measure of initial tempo is shown in Fig. 3b over practice trials for each group. An ANOVA on the mean quarter-note IOIs for the five groups and 11 trials indicated that the mean

[2] We report initial tempo measures rather than average tempo measures because correction errors and pauses, which occurred at different rates over groups, distorted sequence duration measures.

tempo increased (performers played faster) over the five groups ($F(4, 55) = 20.2$, $P < 0.01$). There was also a significant increase in tempo over the eleven practice trials ($F(10, 550) = 18.6$, $P < 0.01$) and a significant interaction between group and practice trial ($F(40, 550) = 2.0$, $P < 0.01$). To address this interaction, separate ANOVAs on trials 1 through 5 and trials 7 through 11 were conducted for each group. Performance by all groups except p1young became significantly faster over the first five trials (Dunn–Bonferroni, $P < 0.05$). Only p3-old and p3-adults showed increased tempo over the last five takes ($P < 0.05$).

We tested whether changes in accuracy and speed over practice followed a power function. Regression analyses predicting both error rates and performance tempo from practice trial were conducted for each performer on both log (base 10) and linear values (log/log and lin/lin), as shown in Table 3. Both functions fit the error rates and tempo values well; however, a power function accounted for more variance than a linear function for all groups for the error rates and for all groups except p1young for the performance tempo. Thus, performance tempo and accuracy improved both with skill level and practice.

4.1.3. Speed/accuracy trade-off

To check whether the changes error rates and initial performance tempo shown in Fig. 3 reflected a speed/accuracy trade-off, the two were correlated across performances: a negative correlation would indicate more errors at faster tempi. In contrast, positive correlations were found between mean initial IOI measures and error rates within performances, both within practice trials across skill levels (ranging from $r = 0.62$ to 0.76, $P < 0.01$) and across practice trials within skill levels (ranging from $r = 0.15$ to 0.55, $P < 0.10$). Although there was limited change in performance tempo overall, the slower performances tended to contain more errors, contrary to predictions of a speed-accuracy trade-off.

4.2. Relative-timing/pitch accuracy trade-off

Some error types reflect failures in accuracy (such as pitch errors), whereas others reflect failures in relative timing (such as duration errors). In this section we examine

Table 3
Regression analyses (r): mean of individual pianists within group ($N = 11$)

	Error rates		Tempo	
	Mean lin/lin	Mean log/log	Mean lin/lin	Mean log/log
p1-young	− 0.72[a]	− 0.79[b]	− 0.76[b]	− 0.71[a]
p2-young	− 0.89[b]	− 0.94[b]	− 0.74[b]	− 0.92[b]
p2-old	− 0.92[b]	− 0.92[b]	− 0.90[b]	− 0.96[b]
p3-old	− 0.84[b]	− 0.95[b]	− 0.92[b]	− 0.98[b]
p3-adult	− 0.53	− 0.59	− 0.83[b]	− 0.99[b]

[a] $P < 0.05$.
[b] $P < 0.01$.

improvement in four error types (pitch, duration, pause, and correction errors) over skill levels and practice trials.

4.2.1. Improvement in each error type over groups and practice trials

The main ANOVA on error rates revealed, in addition to the previous main effects of group and practice trial, a main effect of error type ($F(3, 165) = 31.5$, $P < 0.001$): pauses were the most frequent type of error (50.1% of all errors across groups), pitch and correction errors were intermediate (pitch = 19.5%, corrections = 17.5%) and duration errors were least frequent (12.6%). More important, there was a

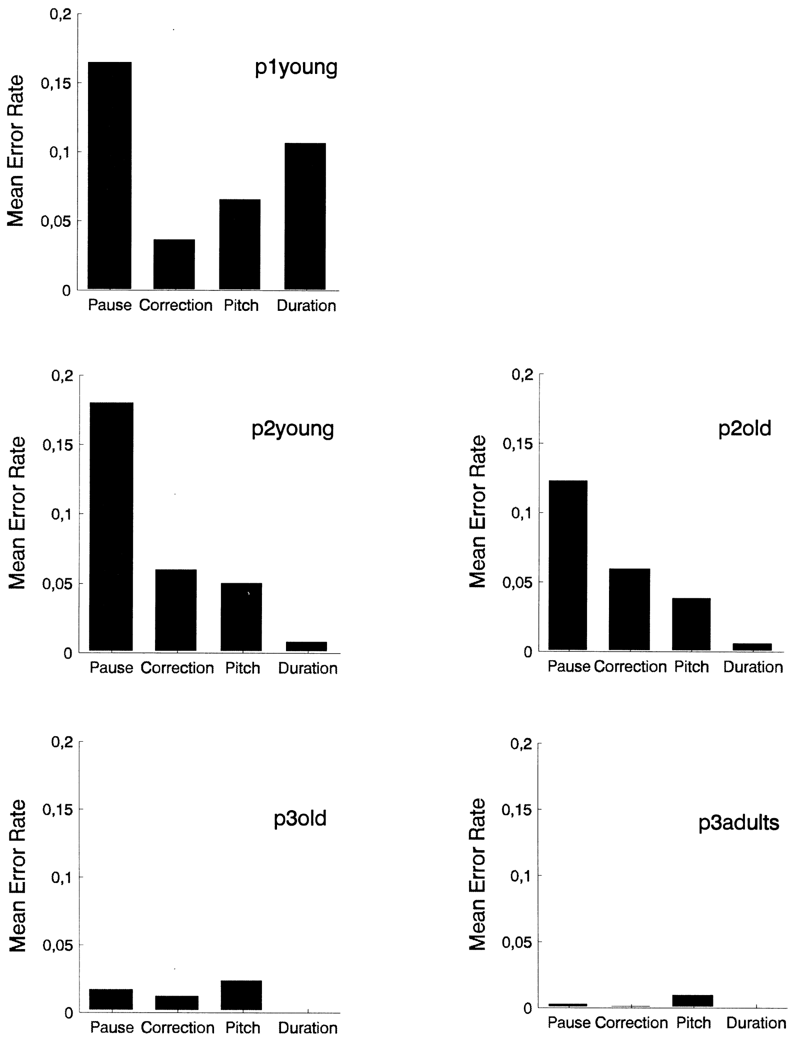

Fig. 4. Mean error rates by error type and group.

two-way interaction between group and error type ($F(12, 165) = 7.9$, $P < 0.01$). As shown in Fig. 4, the least skilled groups (p1-young, p2-young, p2-old) made more pauses, corrections, and duration errors, whereas the most skilled groups (p3-old, p3-adults) made mainly pitch errors. There was also a significant two-way interaction between practice trials and error types ($F(30, 1650) = 5.5$, $P < 0.01$) and a three-way interaction between group, practice trials, and error types ($F(120, 1650) = 1.6$, $P < 0.01$). Although error classifications may limit generalizations about distinctions among error types, the general pattern was that all groups except the adults showed improvements over practice in pauses and corrections (the predominant error types), whereas the adults showed improvements only in pitch error rates (the only predominant error type, as shown in Fig. 4).

4.2.2. Relative timing versus pitch accuracy

We examined the relationship between relative timing and pitch accuracy across skill levels. The percentage of relative timing errors (duration and pause errors) and pitch errors relative to total errors (pitch, duration, pauses and corrections) was computed for each performance, and is shown for each group in Fig. 5a. An ANOVA on these percentages by group, error type (pitch or time), and practice trial (1–11) indicated a significant interaction of group with error type ($F(4, 55) = 18.9$, $P < 0.01$). The less skilled groups (p1-young, p2-young, p2-old) made a higher proportion of timing errors than pitch errors, suggesting that they were concentrating primarily on playing the correct notes irrespective of temporal constraints. The more advanced pianists (p3-old, p3-adults) showed the opposite pattern, with a higher proportion of pitch errors to timing errors. Thus, beginners concentrated primarily on the 'what' to the detriment of the 'when', and advanced pianists mastered the 'when' and only erred on 'what' should be performed.

There was also a significant interaction of error type with practice trial ($F(10, 550) = 2.80$, $P < 0.01$). As shown in Fig. 5b, the proportion of pitch errors increased and timing errors decreased over practice for all groups. Thus, the tendency to err in time shifted to a tendency to err in pitch within the eleven-trial practice session, and across skill levels. An overall negative correlation between pitch and timing error percentages on individual performers' data confirmed this relative-timing/pitch accuracy trade-off ($r = -0.81$, $P < 0.01$).

4.2.3. Tests of independence between pitch and time components

A tradeoff between what pitch events occur and when across the musical sequence suggests that pitch and timing components of individual units also may not be produced independently. To assess their independence, we compute the probability of joint pitch/duration (P/D) errors based on the error rates for the separately occurring pitch (P) errors and duration (D) errors [prob(P) × prob(D)]. As shown in Table 4, P/D errors were observed more often than expected from the individual estimates for all skill groups. For example, the p1young group had twice as many joint P/D errors as expected. Interactions were largest for the most skilled groups; however, even the least skilled group (who produced the most duration errors) showed signif-

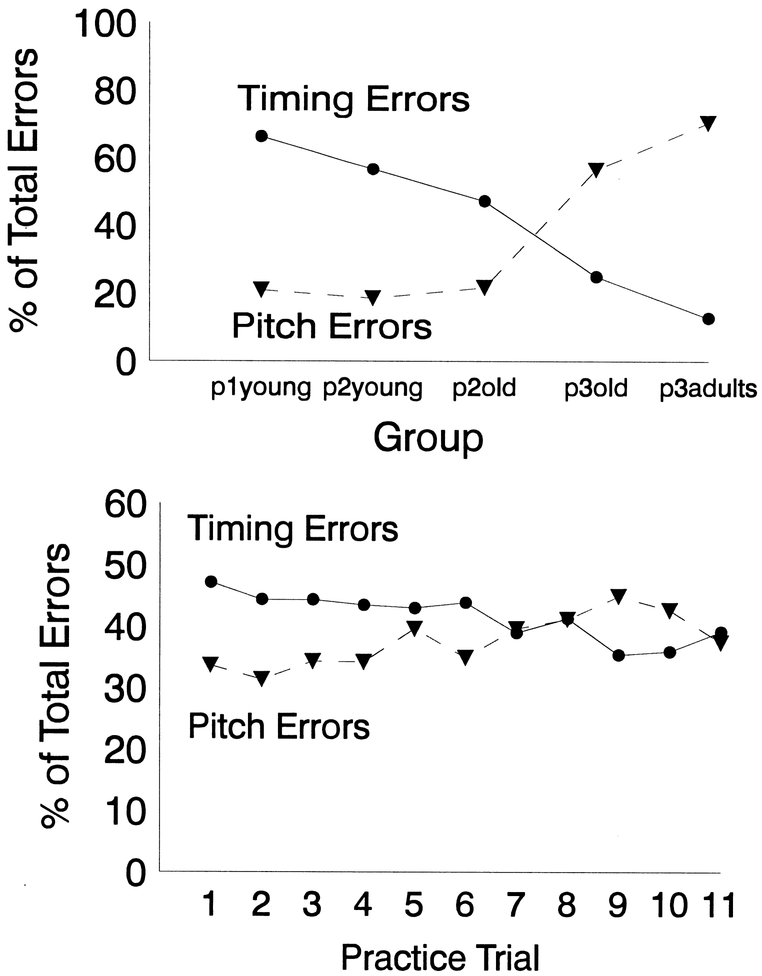

Fig. 5. Percentage pitch and timing errors for the five groups (a) and 11 practice trials (b).

Table 4
Observed and predicted pitch/duration error rates (number per hundred events) and their ratio for the five groups

	Pitch/duration errors (observed)	Pitch/duration errors (predicted) [prob(P) × prob(D)]	Observed/predicted ratio
p1-young	1.48	0.700	2.1
p2-young	0.97	0.033	29.4
p2-old	0.97	0.026	37.3
p3-old	0.19	0.003	63.3
p3-adults	0.04	0.0002	200

icantly more interaction errors than expected. These findings suggest that the pitch and duration components interact even at the individual unit level of performance.

Pitch/duration errors could arise from an attentional failure that affects both dimensions when concentration is lost, or from a memory failure that reflects the retrieval of another unit whose pitch and duration components both differ from those intended. To evaluate these possibilities, we compared whether the pitch and duration components of errors tended to arise from the same source; such an outcome would be consistent with memory retrieval of a different unit whose pitch and duration components were encoded together. Because pitch/duration errors that contained a possible intruder unit (additions and substitutions) formed a small subset of the total errors ($n = 118$), this analysis was conducted over groups. 64.4% of these errors had pitch and duration components that came from the same identifiable source (an event intended for nearby in the sequence), significantly more often than a conservative chance estimate of 50% (binomial test, $P < 0.01$[3]). These findings provide converging evidence that pitch and duration components interact in sequence production, and are probably encoded and retrieved from memory together.

4.3. Mastering temporal structure: temporal continuity, underlying beat and metrical structure

In this section, we further examine errors in music performance as indicators of sensitivity to specific temporal dimensions of music (continuity, beat, and meter).

4.3.1. Temporal continuity

Performers' ability to maintain temporal continuity was measured in terms of the percentage of total errors that disrupted the temporal flow (pauses, corrections, and beat-disruptive duration errors). Three performers (in p3-adults) were excluded from the analysis because they had no pauses, corrections, or beat-disruptive duration errors. An ANOVA performed on the percentage of temporally-disruptive errors by group and practice trial indicated a main effect of group ($F(4, 52) = 20.58$, $P < 0.01$); the ability to respect the temporal flow increased with skill level. There was also a main effect of practice ($F(10, 520) = 3.60$, $P < 0.01$): the ability to respect the temporal flow improved with practice for all groups. Fig. 6 shows the percentages of errors disrupting continuity across groups and practice trials, averaged for the first half of the trials (1–5) and the last half (7–11). There was no interaction of group and practice trial. Thus, performers demonstrated increased ability to maintain temporal continuity with both skill level and specific practice.

4.3.2. Underlying beat

A more specific dimension of temporal structure in music performance is main-

[3] A chance estimate of 50% implies that pitch and duration components have only two possible error outcomes; the musical pieces contained many more pitch and duration alternatives than two, and thus 50% is a conservative chance estimate that nevertheless yields significant results.

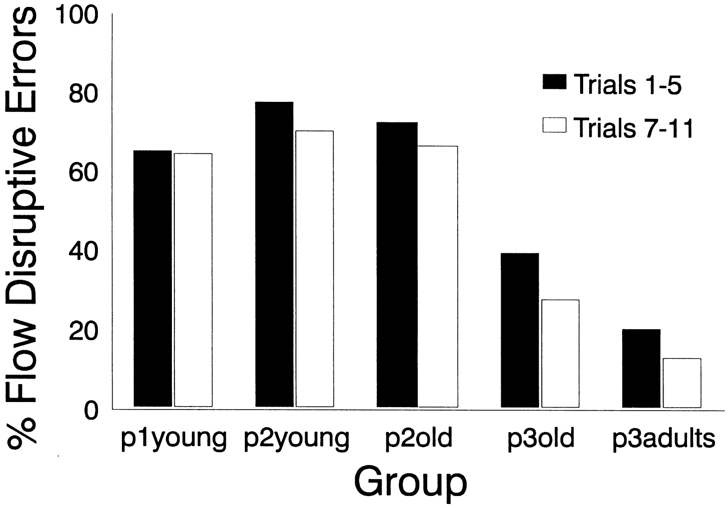

Fig. 6. Percentage flow-disruptive errors for trials 1–5 and 7–11 for the five groups.

tenance of the underlying beat. Some timing errors do not disrupt the musical flow but do disrupt the underlying beat. For example, a quarter-note replaced by a half-note does not disrupt the underlying beat (the quarter-note), but a quarter-note replaced by a single eighth-note does disrupt the underlying beat. The percentage of beat-disruptive errors relative to total number of duration errors was computed for each pianist across trials; there were too few errors to test differences across trials. The percentage of beat-disruptive errors decreased with skill; p1-young = 40.4%, p2-young = 50.3%, p2-old = 35.7%, p3-old = 24.7%, p3-adults = 31.3%. Group data were combined within piece (p2-young and p2-old; p3-old and p3-adults) to provide enough data for group comparisons; beat-disruptive errors were significantly lower than the chance estimate of 50% (probability of the error resulting in an integer multiple of the beat) only for the highest skill group (p3 = 26.5%; $t(14) = -2.71$, $P < 0.05$). Thus, the ability to respect the underlying beat of the music may improve with skill level.

4.3.3. Metrical structure

We tested performers' sensitivity to meter by examining the position of each error type relative to the metrical boundaries indicated by vertical barlines separating measures. Previous evidence from music performance indicated that corrections were more likely to occur at metrical boundaries than at other positions (Palmer & Drake, 1997). ANOVAs on the mean error rates by metrical position and error type indicated a significant interaction between metrical position and error type for each group ($P < 0.05$). Only correction errors indicated effects of metrical position. Therefore, we compared the percentage of total corrections that occurred at metrical barlines for each group, relative to a chance estimate of how often errors were expected to occur at barlines in each piece (number of notated events at metrical

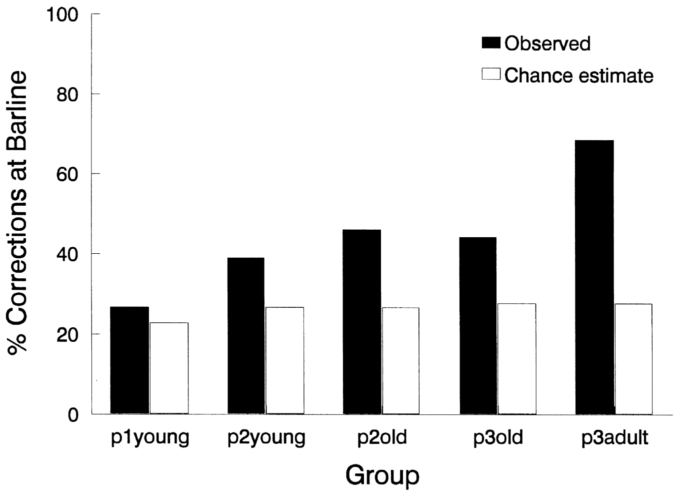

Fig. 7. Predicted and observed percentage corrections at metrical barlines for the five groups.

barlines, relative to total notated events). As shown in Fig. 7, corrections occurred significantly more often at barlines than at any other metrical position for all groups ($P < 0.01$) except p1-young, the least skilled group. Thus, sensitivity to metrical structure was observed early and improved with skill level.

One possible explanation for the increased frequency of corrections at barlines is the physical layout of the musical notation. Performance may be interrupted more often at locations that require large eye movements, such as line endings (which coincide with some metrical barlines). Error percentages and chance estimates for corrections were recomputed, excluding line breaks and events in the metrical bar preceding them. Again, corrections occurred more often at barlines than at other positions for all groups except p1young, relative to chance estimates ($P < 0.01$), similar to the results shown in Fig. 7. Therefore, it was unlikely that the metrical effects were attributable to eye movements alone.

4.4. Preparation of movement: anticipatory behavior and range of planning

Here we examine performers' ability to prepare the contents (pitch) of events for production in terms of future-oriented (anticipatory) behavior and the scope or range of planning.

4.4.1. Anticipatory behavior

We first test differences among skill levels in anticipatory behavior by computing the proportion of anticipatory errors relative to anticipation and perseveration errors combined, for all contextual pitch errors (those in which the intruding event may have been intended for an adjacent position in the sequence). Only errors with an identifiable source in the immediately surrounding context were included; Piece 1 (a single-voiced melody) was not included due to difficulties in coding error movement for single-voice pieces with repeating pitches. An ANOVA on the proportion of anticipatory errors relative to anticipatory plus perseveratory errors (AP) by group (4) indicated a significant effect ($F(3, 44) = 3.4$, $P < 0.05$): anticipatory proportion for p2-young = 50.9%; p2-old = 42.1%; p3old = 62.8%; p3-adults = 62.3%. Thus, the proportion of anticipatory errors increased with skill level, indicating increased future-oriented planning.

We also examined changes in anticipatory proportions across practice. Because these movement errors formed a subset of the overall error data, errors were pooled across groups within trials to examine practice effects. According to Dell et al. (1997) model of sentence production, anticipatory proportion should increase because practice enhances the activation of the present and future relative to the past, and perseverations become less common relative to anticipations as performance improves with practice. The anticipatory proportion increased with practice from the first two trials (AP = 0.55) to the last two trials (AP = 0.68). A regression analysis, predicting AP values by practice trials, indicated that AP increased significantly with practice over the last five trials ($r = 0.91$, $P < 0.05$) and a positive but not significant trend in increased anticipatory behavior over the first five trials ($r = 0.57$, $P > 0.05$). Thus, anticipatory behavior increased with practice.

We next examined factors affecting the relationship between anticipatory proportion and error rate. The Dell et al. (1997) model predicts an inverse relationship between anticipatory proportion and overall error rate (log units), arising from a combination of practice, production rate, activation rate, and decay parameters. Following Dell et al. (1997), we computed the correlation between overall AP and overall pitch error rate (log units) for errors combined across the two Piece 2

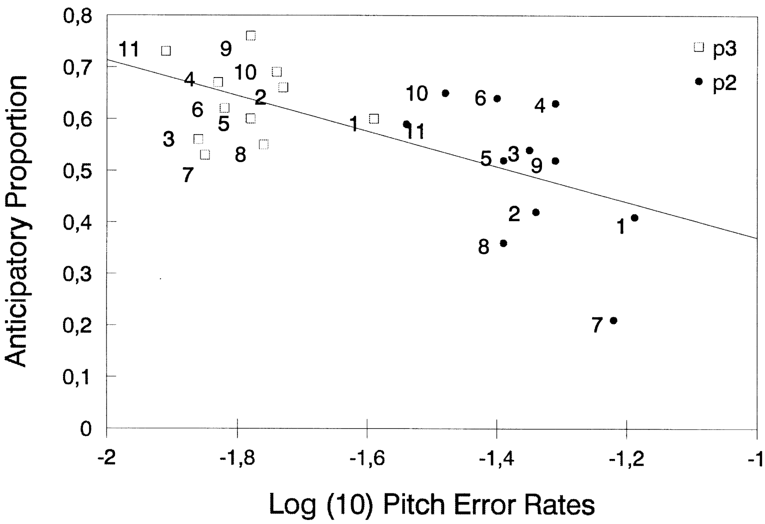

Fig. 8. Regression of total anticipatory proportion on log total pitch error rates for piece 2 and piece 3, over the 11 practice trials.

groups (young and old) and two Piece 3 groups (old and adult), to combine the relatively small error subsets. As shown in Fig. 8, AP increases linearly as log total error rates increase for both groups and practice: $r = -0.64$, $P < 0.01$. Each practice trial within group results in an AP along the regression line, with more advanced performers (p3) and trials representing more practice (11) falling toward the future-oriented (left/high) end of the line. This finding extends the general anticipatory effect (Dell et al., 1997); performances representing higher skill levels and more practice displayed fewer pitch errors overall and more anticipatory behavior.

4.4.2. Range of planning

Another measure of planning ability is the scope or range of planning, reflected in the distance between interacting elements in pitch errors (Palmer & Drake, 1997). The standard deviation of error distances (in number of events) was computed for each group. Only errors with an identifiable source in the surrounding context, defined by the length of the shortest piece (30 events), were included. Bartlett's test of homogeneity of variance revealed significant increases in the standard deviations across the five groups (p1-young = 2.43 events; p2-young = 2.66; p2-old = 3.36; p3-old = 7.04; p3-adults = 7.30; chi-squared (4) = 642.8, $P < 0.01$). Because events can reflect different durations across pieces, error distances across groups were also compared in quarter-note units (the basic beat unit in all of the pieces)[4]. Again, the error distances increased with skill level (p1-young = 2.20 quarter-notes; p2-young = 2.66; p2-old = 3.36; p3-old = 5.91; p3-adults = 6.13; chi-squared (4) = 508.1, $P < 0.01$). In addition, there were significant differences between p2-young and p2-old (chi-squared (1) = 13.3, $P < 0.01$); the older group demonstrated a greater range of planning than the younger group performing the same music. Interacting elements in errors tended to span larger distances as skill level increased, indicating that range of planning increases with skill.

Piece-specific factors, such as pitch repetition rates, may influence the identification of error sources and their distances. Therefore, we computed a chance estimate for each piece based on the mean distance (in number of quarter-notes) between repeating pitches in the musical notation. Analyses conducted only on errors whose distances fell outside the chance estimates confirmed significant differences in standard deviations across groups (p1-young = 4.02; p2-young = 8.62; p2-old = 9.65; p3-old = 9.11; p3-adults = 9.16; chi-squared (4) = 38.6, $P < 0.01$). Thus, the increased range of planning across skill levels was not due solely to differences in piece characteristics.

Finally, we tested whether disruptions in timing of music performance coincide with disruptions in planning capacities. Temporal disruption errors (mean proportions of duration errors and pauses) and range of planning measures (standard deviations of pitch error distances) reflect independent dimensions of performance, as measured by different error characteristics (rates of timing errors and source of

[4] Error distances computed in ms-values across groups did not show an increase with skill level, perhaps due to global tempo differences and rates of temporal disruption across groups, both of which affect durational estimates of error distances.

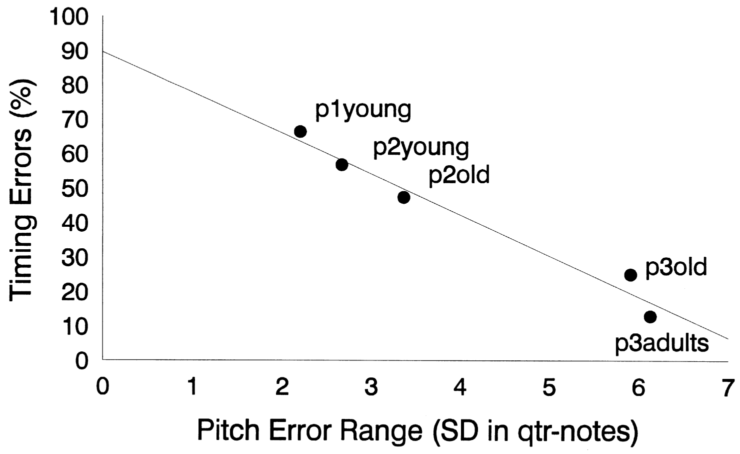

Fig. 9. Correlation between mean timing error percentages and mean range of planning for the five groups.

pitch errors). The two measures were correlated across groups. As shown in Fig. 9, increased skill levels showed both lower proportions of temporal disruptions and a greater range of planning ($r = -0.98, P < 0.01$). As performers' access to sequence events spanned larger ranges, the timing of events was less likely to be disrupted.

5. Summary of findings

5.1. Traditional measures of skill: speed and accuracy

- Both performance tempo and accuracy improved with skill level and practice in early learning trials, consistent with predictions of a power function

.

5.2. Additional characteristic of skill: fluency = relative timing and pitch accuracy

- No speed/accuracy trade-off was observed at any skill level or point in practice.
- A relative-timing/pitch accuracy trade-off was observed, with a shift from focussing on pitch toward mastering time across both skill levels and practice.
- Joint pitch/duration errors occurred more frequently than expected from individual components.

5.3. Mastering the temporal structure of music: temporal continuity, underlying beat and metrical structure

- The ability to maintain temporal continuity improved with skill level and practice.
- The respect of the underlying beat improved with skill level.
- Sensitivity to metrical structure was observed early and improved with skill level.

5.4. Preparation of movement sequences: the ability to plan and the range of planning

- The proportion of anticipatory errors increased with skill level and practice.
- Range of planning (error distances) increased with skill level.
- Range of planning increased as temporal disruptions decreased, with increased skill level.

6. Discussion

This study is the first to rigorously track cognitive changes that result from long-term and short-term learning in music performance. We have documented both quantitative and qualitative differences between expert and novice pianists in how their performances improve with practice. Most studies of music performance focus on one of two issues: systematic expressive nuances that provide perceptual cues to musical structure (e.g. Drake & Palmer, 1993; Palmer, 1989; Penel & Drake, 1998; Repp, 1992a,b; Shaffer et al., 1985; Sloboda, 1983), or production errors that provide cues to planning processes (Palmer & Drake, 1997; Palmer & van de Sande, 1993, 1995; Repp, 1996). Both of these lines of research focus on well-learned performances by highly-trained adult performers, who exhibit fine temporal control and make very few errors. In contrast, the main findings of our study of novice performances indicate that: (1) beginners make many more errors than expert pianists, (2) beginners make different types of errors than experts, (3) beginners exhibit a lack of temporal control, and (4) beginners' ability to plan future events is limited. Therefore, the main theoretical issues that arise in the study of skilled performance are supplanted in novice performance by considerations of relative-timing/pitch accuracy trade-offs, temporal control, and planning constraints.

An original goal of this study was to adapt traditional issues studied in skilled performance to *novice* performance by investigating both expressive nuances and production errors in children's piano performances. However, the highly variable nature of novices' timing made identification of expressive nuances difficult. In addition to the pitch errors described for expert performances, novices' performances exhibited many types of *timing* errors, including corrections, pauses, and duration errors, which are largely absent in adults' well-learned performances (Palmer & Drake, 1997; Palmer & van de Sande, 1995). For example, approximately 25% of the most skilled performers' errors exhibited a timing component in error, whereas more than 75% of the least skilled performers' errors exhibited a timing component in error. Thus, evaluation of novices' performance must include consideration of both which pitches are produced and when they are produced. We have reported a methodology in this study for examining both pitch and temporal errors as indicators of how performers maintain temporal control and plan for future events. In the next section we focus on these findings in terms of the four issues of skill acquisition raised in the introduction.

6.1. Four issues of skill acquisition in music performance

6.1.1. Traditional measures of skill: speed and accuracy

Both performance tempo and accuracy improved across skill levels (demonstrating long-term learning) and with practice (demonstrating short-term learning). More skilled pianists made fewer errors and performed at a faster rate. The observed initial rapid improvement in early practice trials and a gradual slowing in improvement with further practice is consistent with the power law described previously for many perceptual, cognitive, and motor tasks (Newell, 1980; Rosenbaum, 1991). Both the linear and logarithmic fits are within the range of those reported previously (here $r^2 = 0.36$–0.99 compared with $r^2 = 0.40$–0.99 in Newell & Rosenbloom (1981)). In general, a power function fit the performance tempo and accuracy curves better than a linear function, demonstrating that the power function also applies for a complex task such as music performance. Most important, the power function held at initial stages in learning (the first eleven practice trials, compared with hundreds or thousands of trials in previous studies), in contrast to views that a poor fit for early trials reflects increased noise, associated with differences in learning strategies or prior knowledge (Ivry, 1996).

6.1.2. Additional characteristic of skill: fluency = relative timing and pitch accuracy

We did not observe a speed/accuracy trade-off in music performances at any skill level or point in practice, consistent with previous studies of skilled music performances (Palmer & van de Sande, 1993, 1995). This failure may be related to the additional temporal constraints imposed on music performance: in contrast to tasks such as typing, chess and problem solving, in which speed and accuracy are part of the task demands (MacKay, 1971), the task demand in music performance is not to play as fast as possible, but to respect the rhythmic structure specified in the relative timing of sequence events.

We report instead a relative-timing/pitch accuracy trade-off in both novice and expert piano performances for a novel musical sequence: pianists tended to concentrate either on performing the correct pitches or on performing at the correct moment in time. Least skilled pianists' performances exhibited high proportions of timing errors, whereas more skilled pianists exhibited mainly pitch errors. At early levels of skill acquisition, pianists tend to focus preferentially on producing the correct pitches; only later are they able to master the temporal constraints, and thus successfully perform the relative timing and pitch accuracy required of the musical sequence. This relative-timing/pitch accuracy trade-off across skill levels was observed within skill level over practice trials as well; proportionally more timing errors relative to pitch errors were evidenced in initial trials, compared with later trials. Additional support for a relative timing/pitch accuracy trade-off is the frequency of joint pitch/duration errors. They occurred more often, and arose from the same sequence location more often, than expected from individual components. An absence of independence suggests that the pitch and time components intended for each event are accessed together rather than separately. Thus, dimensions of pitch and time interact at the level of individual events as well as at the level of entire sequences. It is possible that stimulus difficulty, which may have changed across groups, contributed to the timing/accuracy trade-off. Difficulty level was only partially controlled in this naturalistic design, and the higher skill levels exhibited lower error rates. However, comparisons of well-learned performances by novice and skilled pianists that contain equivalent error rates confirm that novices tend to produce temporally disruptive errors such as corrections, whereas advanced pianists produce pitch errors almost exclusively (Palmer & Drake, 1997).

6.1.3. Mastering the temporal structure of music: temporal continuity, underlying beat, and metrical structure

Performers' abilities to produce the temporal structure of music can be decomposed into several specific timing skills. One skill is the ability to perform in a continuous manner without pausing or making corrections, which increased across skill levels. Pauses and other forms of temporal disruption have been examined as an index of mental load or production difficulty in complex tasks such as speech (Fodor et al., 1974). Another timing skill is the ability to respect the beat, which also increased across skill levels. A final timing skill is sensitivity to metrical structure, which all groups (except the least skilled) demonstrated: errors occurred more often than chance at the ends of metrical units (barlines) than at other locations, and this ability improved with skill level. These findings are consistent with both performance studies (Palmer, 1989; Palmer & Drake, 1997) and perceptual studies (Palmer & Krumhansl, 1990; Yee, Holleran & Jones, 1994) that demonstrate musicians' increased sensitivity to additional hierarchical levels of temporal structure. The ability to maintain temporal continuity improved with both short- and long-term learning, but improvements in the underlying beat and sensitivity to metrical structure have been demonstrated only over (long-term) learning. Findings that incorporate larger error data sets may also reveal short-term improvements.

6.1.4. Preparation of movements: the ability to plan and the range of planning

Another hallmark of performance is the ability to prepare actions in advance of their production, or to plan. Future-oriented planning was indexed in music performance by the proportion of anticipatory errors (those reflecting influences of events intended for future events) relative to perseveratory errors (those reflecting the continuing influence of events that have already been performed). Our findings that anticipatory behavior increased both with skill level and with practice extend previous findings in music performance and speech production (Dell et al., 1997; Palmer & Drake, 1997) to indicate that experience with a task reflects more future-oriented behavior and less emphasis on the past. Anticipatory and perseveratory errors arise in associative models of planning in sentence production from activation rising more quickly for unintended sequence events than for intended (correct) elements, due to priming from other related events or nodes (Dell, 1986; MacKay, 1987).

The Dell et al. (1997) model of sentence production predicts an increased anticipatory proportion of errors as practice enhances the activation of events intended for the present and future relative to the past. In addition, their model predicts an inverse relationship between anticipatory proportion and overall error rate, arising from a combination of practice, production rate, activation rate, and decay parameters. Our findings indicate that musicians were indeed more likely to anticipate as practice increased, and the anticipatory proportion was inversely related to overall error rates, both across skill levels and practice. The inverse relationship between error rates and anticipatory behavior, referred to as a general anticipatory effect, extends the speech error findings to a task that entails strict temporal constraints (Dell et al., 1997). These similarities suggest that serial order processes in sequence production reflect similar planning abilities at least across domains of music and speech.

A related indicator of planning is the ability to plan for events from larger ranges of the sequence, referred to as range of planning. Errors tended to reflect events from increasingly larger ranges of the sequence at advanced skill levels, similar to previous findings with well-learned music performances (Palmer & Drake, 1997). In fact, the average number of events spanned by errors in performances of unfamiliar music (two to seven events) was less than the typical range seen in well-learned performances (six to nine events) by pianists of similar skill levels (beginners, intermediates, and advanced pianists). Thus, this measure of planning scope suggests that the availability of elements from longer sequence ranges increases with both short-term practice with a piece and long-term experience. Furthermore, range of planning abilities are related to temporal control. There was an inverse relationship across skill level between performers' (pitch) range of planning and the frequency of temporal disruptions. These findings suggest similar underlying cognitive abilities for temporal control and planning processes that develop with skill.

6.2. Integrating hallmarks of music performance: temporal control as an indicator of planning constraints

Two noteworthy improvements in music performance over both short-term and long-term learning are highlighted in this study, namely, the progressive mastering of temporal constraints and increased planning abilities. The strong relation between these abilities suggests that the two phenomena are reflections of the same underlying cognitive processes. We first identify functional characteristics of performance planning, and then suggest how breakdowns in temporal control reflect this planning.

6.2.1. The planning process

Music performance in a sight-reading task involves the identification of musical events, the conceptual interpretation of their structural relationships, and the co-ordination of movements necessary to produce them. Most studies of (adult) skilled performance evaluate the final product: the often-memorized production of a well-learned musical piece. These studies usually indicate an error-free, smooth, uninterrupted musical flow. In contrast, performances of novel musical pieces often lead to errorful performances that are temporally irregular (especially for novices). Such breakdowns in performance provide an opportunity to identify certain functional characteristics of the cognitive abilities underlying performance, under less-than-ideal conditions. Based on these observations and related findings, we identify the following functional characteristics:

- *Segmentation of a sequence*. A long musical sequence is not planned all at once, but instead is broken down into short segments that contain several events (van Galen & Wing, 1984). Elements within a segment are more likely to be planned at the same time because they are simultaneously accessible; as a result, the relative timing of events within a segment tends to be respected, and events within a segment are more likely to interact with each other in errors.
- *Size of segments.* Segment size is determined primarily by perceptual and memory constraints that limit the number of accessible events at a given time during the performance. However, segments may not reach the maximal planning capacity in sight-reading tasks because structural relationships (such as phrase boundaries) and physical relationships (such as note spacing) can lead to the creation of segment boundaries before the maximal planning capacity is reached (Sloboda, 1977).
- *Hierarchical structure*. Planning over longer time-spans is facilitated if the sequence is conceived in a hierarchical fashion. Temporally remote, non-adjacent events that are related at higher hierarchical levels, such as metrical levels, as may be conceived and planned simultaneously. Relative timing can be respected over a longer time-spans than individual segments, and increasingly remote events become simultaneously accessible in a larger range of planning as hierarchical relationships are learned.
- *Planning near segment boundaries*. Although planning occurs continuously during performance, planning for a future segment is concentrated near the end of the current segment. Indicators of breakdowns in planning (such as pauses or corrections) are more likely to occur at positions at which demands on planning are greatest: near the ends of segments.
- *Simultaneous planning and execution*. Because music performance tasks contain strong temporal demands (not allowing a start-and-stop strategy), events must be executed at the same time as future events are being planned, within the available processing time (determined by temporal factors such as production rate and intended event duration). Breakdowns in performance are more likely to occur when there is insufficient processing time.

Certain predictions for skill acquisition processes arise from these cognitive abilities. First, skilled performers develop an increased range of planning, which arises from both increased planning capacities and heightened sensitivity to hierarchical structures relating non-adjacent sequence elements. Second, skilled individuals develop the ability to anticipate; this ability increases both with long-term and short-term learning. Third, requirements on processing time are reduced with increased skill, leading to greater ability to plan and execute simultaneously, as reflected in fewer temporal disruptions. Finally, increased segment size leads to fewer segment boundaries. Because temporal disruptions occur more frequently when demands of planning the next segment increase (near segment boundaries such as metrical boundaries), the number of opportunities for temporal disruptions decreases as the range of planning increases, leading to an inverse relationship. The findings reported here are consistent with each of these predictions.

6.2.2. Relation between planning and temporal control

Under the strict temporal constraints imposed in music performance, performers must produce the correct events (what) *and* produce them at the correct moment in time (when). Production errors, or breakdowns in planning, arise when performers cannot produce the intended events within the time available. When this happens, two possible scenarios arise: the first is to focus on the event contents and correctly produce the intended pitches, at whatever time they become available. Another is to focus on the time at which the contents must be produced, and the content reflects whatever events are accessible at the correct time.

The *novice* performances described in this study typify the first solution; their performances are frequently interrupted by pauses, corrections, and duration errors. Additional processing time is obtained simply by stopping: future events are performed only after planning has been completed. Because segments are short (as evidenced by small ranges of planning), there are frequent opportunities for temporal breakdown at boundaries. The *skilled* performances typify the second solution. The correct pitches and relative timing can be planned efficiently in the available processing time because skilled performers require less processing time, and so performance is accurate in relation to both pitch and time. Moreover, planning segments are long, resulting in fewer opportunities for temporal disruption. Skilled performances of difficult tasks (such as sight-reading, or performing at a fast tempo) still reflect the second solution; inadequate processing time results in sub-optimal planning, leading to temporally correct performances that contain more pitch errors. Thus, in skilled performance, breakdowns in planning are primarily in pitch, and rarely in time.

Although the influence of timing demands on planning constraints (and vice versa) may not come as a surprise to researchers of music performance, considerations of timing do not enter into many theories of skilled behaviors. This may be attributable in part to the fact that most research examines well-learned behaviors, or behaviors of highly skilled individuals, who may exhibit very little temporal disruption. However, the greatest cognitive change in sequence learning may be found at relatively inexperienced skill levels, and so novice performance may be an important indicator of the cognitive bases of skill acquisition (Adams, 1987; Palmer &

Drake, 1997; Palmer & Meyer, in press). That finding, combined with the parallels reported here in planning behaviors that transcend music and speech, suggest that examination of sequence timing may explicate planning abilities that underlie many complex skills.

Acknowledgements

This research was supported in part by NIMH Grant R01-45764 to the second author. We are grateful for assistance and comments from Melissa Blakeslee, Gary Dell, Kory Klein, Rosalee Meyer, Pete Pfordresher, Bruno Repp, Grant Rich, Brent Stansfield, and Tim Walker.

References

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin*, *101*, 41–74.

Anderson, J. R. (1982). Acquisition of cognitive skills. *Psychological Review*, *89*, 369–406.

Bamberger, J. (1980). Cognitive structures in the apprehension and description of simple rhythms. *Archives de Psychologie*, *48*, 171–199.

Chi, M.T.H., Glaser, R., Farr, M.J. (1988). The nature of expertise. Hillsdale, NJ: Erlbaum.

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

Cooper, G., & Meyer, L. B. (1960). *The rhythmic structure of music*, Chicago: The University of Chicago Press.

Crossman, E. R. F. W. (1959). A theory of the acquisition of speed skill. *Ergonomics*, *2*, 153–166.

Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: a functional analysis and a model. *Psychological Review*, *104* (1), 123–147.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93* (3), 283–321.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition*, Orlando: Academic Press.

Drake, C., & Gérard, C. (1989). A psychological pulse train: how young children use this cognitive framework to structure simple rhythms. *Psychological Research*, *51*, 16–22.

Drake, C., & Palmer, C. (1993). Accent structures in music performance. *Music Perception*, *10* (3), 343–378.

Drake, C., Dowling, J., & Palmer, C. (1991). Accent structures in the reproduction of simple tunes by children and adult pianists. *Music Perception*, *8* (3), 313–332.

Drake, C. (1993a). Reproduction of musical rhythms by children, adult musicians and adult non-musicians. *Perception and Psychophysics*, *53* (1), 25–33.

Drake, C. (1993b). Influence of age and experience on timing and intensity variations in the reproduction of short musical rhythms. *Psychologica Belgica*, *33-2*, 217–228.

Drake, C. (1997). Motor and perceptually preferred synchronisation by children and adults: binary and ternary ratios. *Polish Quarterly of Developmental Psychology*, *3* (1), 41 –59.

Drake, C. (1998). Psychological processes involved in the temporal organization of complex auditory sequences: universal and acquired processes. *Music Perception*, *16* (1), 11–26.

Ericsson, K. A., & Staszewski, J. (1989). Skilled memory and expertise: mechanisms of exceptional performance. In D. Klahr, & K. Kotovsky, *Complex information processing: the impact of Herbert A. Simon*, (pp. 235–367). Hillsdale, NJ: Erlbaum.

Ericsson, K. A., & Smith, J. (1991). Prospects and limits of the empirical study of expertise: an introduction. In K. A. Ericsson, & J. Smith, *Toward a general theory of: prospects and limits*, (pp. 1–39). Cambridge: Cambridge University Press.

Ericsson, K. A., Krampe, R., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363–406.

Essens, P. J., & Povel, D. J. (1985). Metrical and non-metrical representations of temporal patterns. *Perception and Psychophysics*, *37* (1), 1–7.

Essens, P. J. (1986). Hierarchical organization of temporal patterns. *Perception and Psychophysics*, *40*, 69–73.

Essens, P. J. (1995). Structuring temporal sequences: comparison of models and factors of complexity. *Perception and Psychophysics*, *57* (4), 519–532.

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language: an introduction to psycholiguistics and generative grammar*, New York: McGraw-Hill.

Garcia-Albea, J. E., del Viso, S., & Igoa, J. M. (1989). Movement errors and levels of processing in sentence production. *Journal of Psycholinguistic Research*, *18*, 145–161.

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower, *Psychology of learning and motivation*, (pp. 133–177). New York: Academic Press.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth, *Language production: Speech and Talk*, (pp. 177–220). London: Academic Press.

Gruson, L. M. (1988). Rehearsal skill and musical competence: does practice make perfect? In J. A. Sloboda, *Generative processes in music: the psychology of performance, improvization, and composition*, (pp. 91–112). New York: Oxford University Press.

Ivry, R. (1996). Representational issues in motor learning: phenomena and theory. In H. Heuer, & S. W. Keele, *Handbook of perception and action, vol 2: motor skills*, (pp. 263–330). San Diego, CA: Academic Press.

Keele, G. E., & Summers, J. (1976). The structure of motor programs. In Stelmach, *Motor control: issues and trends*, (pp. 109–142). New York: Academic Press.

Kolers, P. A. (1975). Memorial consequences of automatized encoding. *Journal of Experimental Psychology: Human Learning and Memory*, *1* (6), 689–701.

Lehmann, A. C., & Ericsson, K. A. (1993). Sight-reading ability of expert pianists in the context of piano accompanying. *Psychomusicology*, *12*, 182–195.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*, Cambridge Massachusetts: MIT Press.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527.

MacKay, D. (1971). Stress pre-entry in motor systems. *American Journal of Psychology*, *84*, 31–35.

MacKay, D. G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review*, *89*, 483–506.

MacKay, D. G. (1987). *The organization of perception and action: a theory for language and other cognitive skills*, New York: Springer.

McKenzie, C. L., Nelson-Schultz, J. A., & Wills, B. L. (1983). A preliminary investigation of motor programming in piano performance as a function of skill level. In D. R. Rogers, & J. A. Sloboda, *The acquisition of symbolic skills*, (pp. 283–292). New York: Plenum Press.

Newell, A. (1980). Harpy, production systems and human cognition. In R. Cole, *Perception and production of fluent speech*, (pp. 289–380). Hillsdale, N.J: Erlbaum.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson, *Cognitive skills and their acquisition*, (pp. 1–55). Hillsdale, NJ: Erlbaum.

Palmer, C., & Drake, C. (1997). Monitoring and planning capacities in the acquisition of music performance skills. *Canadian Journal of Experimental Psychology*, *51* (4), 369–384.

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, *16* (4), 728–741.

Palmer, C., & Meyer, R. K. (in press). Conceptual and motor learning in music performance. *Psychological Science*,.

Palmer, C., & van de Sande, C. (1993). Units of knowledge in music performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19* (2), 457–470.

Palmer, C., & van de Sande, C. (1995). Range of planning in skilled music performance. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 947–962.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance*, *15* (12), 331–346.

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

Parncutt, R. (1993). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, *11* (4), 409–464.

Penel, A., & Drake, C. (1998). Sources of timing variations in music performance: a psychological segmentation model. *Psychological Research*, *61* (1), 12–32.

Povel, D. J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Cognition*, *7*, 3–18.

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

Repp, B. H. (1996). The art of inaccuracy: why pianists' errors are difficult to hear. *Music Perception*, *14*, 161–183.

Rosenbaum, D. A. (1991). *Human motor control*, San Diego: Academic Press.

Schneider, W. S. R. (1977). *Controlled and automatic human information processing: I. detection, search, and attention. Psychological Review*, *84* (1), 1–66.

Shaffer, L. H. (1980). Analysing piano performance: a study of concert pianists. In G. E. Stelmach, & J. Requin, *Tutorials on motor behavior*, (pp. 443–455). Amsterdam: North-Holland.

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

Sloboda, J. A. (1977). Phrase units as determinants of visual processing in music reading. *British Journal of Psychology*, *68*, 117–124.

Sloboda, J. A. (1983). The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology*, *35*, 377–396.

Smith, J. (1983). Reproduction and representation of musical rhythm: the effects of musical skill. In D. R. Rogers, & J. A. Sloboda, *Acquisition of symbolic skill*, (pp. 273–283). New York: Plenum.

van Galen, G., & Wing, A. M. (1984). The sequencing of movements. In M. Smyth, & A. M. Wing, *The psychology of human movement*, (pp. 153–182). London: Academic Press.

Yee, W., Holleran, S., & Jones, M. (1994). Sensitivity to event timing in regular and irregular sequence: influences of musical skill. *Perception and Psychophysics*, *56*, 461–471.

Yeston, M. (1976). *The stratification of musical rhythm*, New Haven: Yale University Press.