Music Perception
Winter1996, Vol. 14, No. 2, 161–184

The Art of Inaccuracy: Why Pianists' Errors Are Difficult to Hear

BRUNO H. REPP
Haskins Laboratories, New Haven, Connecticut

Pianists' pitch errors were identified in a MIDI data base comprising more than 90,000 notes. Ten graduate student pianists had played four pieces (Schumann's *Träumerei*, Debussy's *La fille aux cheveux de lin*, Chopin's Prelude in D-flat Major, and Grieg's *Erotik*) three times from the score, after only a brief rehearsal. Pitch errors were classified exhaustively as substitutions, omissions, or intrusions. (A frequent form of intrusion was the "untying" of tied notes.) Nearly all errors occurred in nonmelody voices, often inside chords. The majority of the intrusions and nearly all substitutions seemed contextually appropriate. The repeated performances made it possible to distinguish consistent from unique errors. Consistent errors were more often omissions than intrusions, and consistent intrusions were more contextually appropriate than unique intrusions. Most errors seemed likely to be perceptually inconspicuous. This was confirmed in an error-detection experiment, in which eight pianists, some of whom had recently studied the test piece (the Chopin prelude), collectively detected only 38% of all objectively registered errors. Pitch errors, rather than being a categorical phenomenon (as a score-based analysis might suggest), vary in the degree to which they violate the music, and their perceptibility is context-, listener-, and situation-dependent. Members of a typical concert audience are likely to notice only a small fraction of a pianist's inaccuracies, which is in part due to the contextual appropriateness of most errors.

Introduction

Even in this age of technical proficiency, live musical performances are rarely perfect, especially below the highest levels of accomplishment. Lack of concentration, technical deficiencies, insufficient rehearsal time, poor sightreading ability (in "quick study" situations), and a willingness to take technical risks for the sake of expression are among the factors that conspire to cause errors of various kinds and degrees of severity. Although such imperfections can be a source of considerable chagrin for the musi-

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

cians, their audience tends to be much less aware of them. Sloboda (1985) observed wisely that "[e]xperienced performers soon come to realize just how much they can 'get away with' in live performance. I have often been amazed, when listening to a recording of my own performance, just how unnoticeable were errors which, at the time of performance, struck me as catastrophic. Indeed, part of the art of sight reading is knowing which parts of the music will not be salient for a listener. One learns how to create an impression of accuracy in a performance that is actually far from faithful to the score" (p. 85).

"Faithfulness to the score" is the generally accepted criterion for deciding what constitutes an error in performance. Broadly understood, that concept invokes a host of complex issues that are far beyond the scope of the present paper (see Taruskin, 1995). This study deals only with a limited class of errors, pitch errors in piano performance of tonal "classical" music, whose definition is straightforward—or so it seems. There are three types of such errors: substitutions, omissions, and intrusions.[1] A substitution is the playing of a note with the wrong pitch, such as E_4 instead of C_4. The underlying assumption (which seems justified in most instances) is that the pianist either misread C_4 as E_4 or intended to play C_4 but hit E_4 instead. Occasionally, however, such an error may arise from the simultaneous but independent occurrence of an omission and an intrusion in the same chord. An omission is the failure to play a note that is in the score, whereas an intrusion is the playing of a note that is not in the score. A special kind of intrusion is the "untied note," which does appear in the score but is tied to a previous note of the same pitch and thus is not intended to be sounded again.

Performance involves a number of different stages or levels at which errors may be observed: The score (symbolic level) is read by the pianist (level of visual perceptual and cognition), who moves arms and fingers (kinematic level) to depress keys that hurl hammers against strings (mechanical level), which results in tones (acoustic level) that are heard by listeners, including the pianist (level of auditory perception and cognition). In the score, pitches are represented by the vertical placement of note heads that are either present or absent.[2] In the actual music, however, pitches are conveyed by tones that have particular durations and intensities and occur

1. Omissions are also called "deletions," and intrusions are also called "additions" (Palmer & van de Sande, 1993). However, these latter terms have an undesirable connotation of intentionality that the present terms avoid. Palmer and van de Sande also assume that pitches, like phonemes in speech errors, can "move" from one location to another, which leads to additional error categories ("shifts" and "exchanges"). No such process assumptions are made here, and the three error categories are therefore exhaustive.

2. For the sake of this discussion, we may assume a one-to-one correspondence of notes, tones, and perceived pitches (fundamental frequencies). This is not to deny that it is occasionally possible to perceive pitches that are not notated in the score, such as the common fundamental (root) of two or more simultaneous notes.

in certain contexts. The presence or absence of these tones in the musical fabric is no longer a categorical matter. For example, a tone may be too faint in its context to be heard by a listener, perhaps even too faint (e.g., because of masking by other tones) to be detected in an objective acoustic analysis. This may be equally true for a correctly played note and for an intrusion error. Is the intrusion then an error? And if the inaudible correct note were omitted, would this be an omission error? The answer depends on the level of analysis. At the level of a listener's perception, the answer would have to be negative. At the mechanical level, however, one can register objectively that the pianist depressed an incorrect key (and achieved hammer-string contact) or failed to depress a correct key, regardless of the perceptibility of the consequences. At that level, pitch errors (really: key-depression errors) can be defined objectively and unambiguously.

The definition of pitch errors in piano performance thus depends on the level that is at the focus of attention. Listeners operate mainly on the perceptual level; pianists focus both on the kinematic and perceptual levels, perhaps giving emphasis to one or the other in different situations; but psychologists who investigate music performance tend to focus on the mechanical level, because of the objectivity it affords and also because of its accessibility through the Musical Instrument Digital Interface (MIDI) systems that are now widely available. However, it is at the perceptual level that it is decided whether an error really "counts." An error that is registered objectively via MIDI is of little musical and aesthetic significance if it is not detected by most listeners. More precisely, it is of no significance at all for those listeners who do not hear it, and its significance for those listeners who do hear it depends on its nature and degree of severity. It seems plausible that more errors will be registered objectively (via MIDI) than perceptually, although the magnitude of the discrepancy is not known. The present study is concerned with this discrepancy and its causes.

It will be possible to address only some aspects of this issue here. The detectability of any objectively registered pitch error in a piano performance depends on three factors: the auditory prominence of the tone in its context, the listener's auditory sensitivity and musical experience, and the listening situation. Only the first factor—itself a complex of variables—will be considered here in detail, and then only in terms of plausible conjecture at the level of MIDI data, corroborated by a small eror-detection experiment. The other two factors can only be acknowledged here. Obviously, the detectability of all but the most egregious errors depends on whether the listener is musically trained and knows the score. Most likely, it also depends on whether the listener is looking at the score, whether the performance is presented just once or repeatedly or in small segments, and whether the location of potential errors is known in advance. The error-detection experiment reported in this paper represents just one of many possible scenarios: Undergraduate pianists familiar with the music listened to uninter-

rupted performances with the (unmarked) score before them. The intention was to stay reasonably close to musical practice, not to conduct a psychophysical experiment.

The purpose of the error-detection experiment was to confirm the author's informal observation that many MIDI-registered errors are difficult to hear, at least in the context of an uninterrupted performance, and to examine what differentiates errors that are easy to detect from those that pass by unnoticed. The errors occurred in performances recorded as part of a MIDI data base for the study of expression and were mainly due to the pianists' limited rehearsal time. As a "field study" of the nature and perceptibility of such "uninvited" errors, the present research complements recent laboratory studies in which errors were elicited or at least welcomed (Palmer & van de Sande, 1993, 1995) or artificially introduced (Palmer & Holleran, 1994). The error-detection experiment was preceded by an analysis of all pitch errors in the MIDI corpus. This analysis was less an attempt to learn about cognitive processes underlying music performance (which are better studied in the laboratory, as Palmer & van de Sande did) than it was perceptually motivated: It was assumed (cf. Sloboda's observations cited earlier) that skilled pianists avoid very obvious mistakes and therefore mostly commit errors that are difficult to hear. Thus, specific hypotheses about factors that determine the relative perceptibility of errors were simultaneously hypotheses about the nature and distribution of these errors in the music. The following hypotheses were considered:

1. Errors will be less noticeable (and will occur more frequently) in subsidiary voices, especially inner voices, than in the principal melody (which is most often the highest voice). Both predictions were confirmed in studies by Palmer and colleagues, who found that errors in an inner voice were most difficult to detect (Palmer & Holleran, 1994) and that errors were less likely to occur in the melody voice than in nonmelody voices, and were also less likely in the highest-pitched voice, regardless of whether or not it was considered the melody (Palmer & van de Sande, 1993).
2. Errors will be less noticeable (and will occur more frequently) when other note onsets are present at the same time, in proportion to the number of simultaneous onsets.[3] A relevant finding is that voice entries (especially inner voice entries) in polyphonic music are difficult to hear when the number of voices increases (Huron, 1989).[4]

3. An increase in the probability of errors with the number of notes in a chord is predicted on purely statistical grounds.

4. A corollary of this hypothesis is that intrusion and substitution errors will be more noticeable when they are asynchronous with other simultaneous notes, especially when their onset precedes that of the other tones (cf. Rasch, 1978). However, the temporal alignment of tones was not analyzed in detail in the present study.

3. Intrusion and substitution errors will be less noticeable when they fit with their context, melodically and/or harmonically, than when they clash with it. Correspondingly, it was predicted that the majority of registered errors will be contextually appropriate. Palmer and van de Sande (1993) found support for the latter prediction in homophonic music, but the proportion of contextually appropriate errors was lower in polyphonic music.
4. One special feature of the present data base was that it contained repeated performances of each piece by the same pianists. This made it possible to distinguish consistent (twice repeated) from partially consistent (once repeated) and unique errors. In addition, errors committed by several pianists could be distinguished from idiosyncratic errors. It stands to reason that pianists will be more likely to repeat errors they themselves have not noticed. It was thus predicted that contextually appropriate errors will be more consistent than inappropriate errors and that, conversely, consistent errors will tend to be contextually appropriate.
5. Intrusion and substitution errors will be less noticeable when they are relatively low in intensity (MIDI velocity). Correspondingly, intrusion and substitution errors may tend to be softer than correct notes. This hypothesis was investigated only in the performances that were used in the error-detection experiment.

Other possible hypotheses could be envisioned. For example, one might suppose that errors are less noticeable and/or more frequent in metrically weak than in metrically strong positions, or that the presence of notes of the same pitch nearby inhibits detection and/or promotes occurrence of errors. Palmer and van de Sande (1993) noted that pitch errors often occur in the vicinity of notes of the same pitch and interpreted these errors as anticipations or perseverations. Later, they showed that these errors occur mainly within phrases and that error patterns change accordingly when pianists are instructed to phrase the same music in different ways (Palmer & van de Sande, 1995). These are interesting findings, but there may be other possible causes of such errors, such as technical simplification of a passage involving repeated notes. Musical compositions have many special features that may lead to highly context-specific patterns of errors, often due to the technical (and, sometimes, reading) problems posed by particular constellations of notes. An understanding of these problems requires that errors be considered in their specific contexts; their causes usually can only be guessed at. Therefore, hypotheses about metrical structure or sequential pitch relationships are better addressed with controlled materials in the laboratory than in a "field study" of expressive performances. However, some of the observations to follow may be relevant to these hypotheses.

Methods

THE MUSIC

Since the purpose of collecting the present MIDI data base was to study expression, not errors, pieces had been selected that were musically interesting but technically not very demanding. They all had a moderate or slow tempo and required much legato articulation and pedaling. They were *Träumerei* (No. 7 of the suite *Kinderszenen*, op. 15) by Robert Schumann (Breitkopf edition); *La fille aux cheveux de lin* (No. 8 of the *Preludes*, Book I) by Claude Debussy (Durand edition); Prelude in D-flat Major (No. 15 of the *24 Preludes*, op. 28) by Frédéric Chopin (Schott/Universal edition); and *Erotik* (No. 5 in Book III of the *Lyric Pieces*, op. 43) by Edvard Grieg (Peters edition). The first three pieces were expected to be familiar to all participating pianists, at least from listening; the Grieg piece, however, was deliberately selected to be less familiar.

The score of the Chopin prelude, which was used in the error-detection experiment, is included as an appendix. Space limitations preclude reproduction of the other musical scores here, although it will occasionally be necessary to refer to them. Specific positions in the music will be denoted by the convention "bar-beat-subdivision"; thus, "15-3-2" refers to the note(s) on the second half-beat in the third beat of bar 15.

THE PIANISTS

Ten pianists (P1, P2, ..., P10) participated in the study. Nine of them were graduate students of piano performance at the Yale School of Music: five first-year, one second-year, and three third-year (artist's diploma) students. One pianist was about to graduate from Yale College and had been accepted by the School of Music. The pianists' age range was 21 to 29 years, and they had started to play the piano between the ages of 4 and 8 years. Seven were female, three male.

PROCEDURE

The recording took place in a large room housing an upright Yamaha MX100A Disklavier connected to a Macintosh computer. The pianist was given the music and had 1 hr to rehearse the four pieces. Subsequently, the pieces were recorded once, in whichever order the pianist preferred, and then two more times in the same order. The pianists played from the score and were asked to give special attention to expression. If something went seriously wrong in a performance, it was repeated immediately. One pianist, P4, was able to record only two performances of each piece; all others recorded three, as planned. At the end of the session, each pianist filled out a questionnaire and was paid $50.

The questionnaire inquired in some detail about the familiarity of the pieces. Schumann's *Träumerei* had been previously studied (at any time in the past) by three pianists and played informally by two; the rest knew it well from listening only. Debussy's *La fille aux cheveux de lin* had been studied by three pianists, played informally by three, and heard repeatedly by four. Chopin's Prelude in D-flat Major had been studied by four, played informally by three, and heard by three. Grieg's *Erotik* was totally unfamiliar to 9 of the 10 pianists; only one of them had studied it. With few exceptions, therefore, the performances could be characterized as "prepared sightreading" or "quick study." The pianists were also asked to indicate how satisfied they were with their performances, choosing from the categories "best effort," "good effort," "average," "below average," and "poor." All pianists chose the middle three categories, with the following frequencies: 4, 5, 1 for Schumann; 6, 4, 0 for Debussy; 3, 4, 3 for Chopin; and 4, 5, 1 for Grieg.

DATA ANALYSIS

The MIDI data were imported as text files into a spreadsheet program where the note onsets were extracted, labeled, and compared with a numerical (MIDI pitch) transcription

of the score. In that laborious process (there were 93,264 notes to be verified), all pitch errors were identified and labeled. The errors were counted and classified for the present analyses. Rather than quantifying the contextual appropriateness of intrusion and substitution errors by some objective method (which would have been prohibitively time-consuming and of uncertain perceptual relevance), the author made rough perceptual judgments by playing the chords containing the wrong notes on the piano and deciding whether they sounded harmonious or jarring. (Harmonic, rather than melodic, appropriateness was at issue in nearly all instances.) Although this method obviously involved a subjective criterion of dissonance, it seemed adequate for the present purpose, which was to predict which errors would be easy to hear and which difficult—predictions that were tested, for a subset of the data, in the error-detection experiment.

Results and Discussion

ERROR FREQUENCIES

Table 1 summarizes the frequencies and percentages of the different types of errors in each piece. It also lists the total number of notes played for each piece. It is evident that omissions were more frequent than intrusions (not including "untied" notes in the Schumann and Debussy pieces), which in turn were more frequent than substitutions; this was true in each piece, although the relative frequencies varied. Untied notes were much more frequent in the Schumann piece, which contains 49 tied notes, than in the Debussy, which contains 71. (The Chopin and Grieg pieces contain hardly any tied notes.) The pianists' lack of familiarity with Grieg's *Erotik* is reflected in the high omission and intrusion rates for that piece. (Its technical difficulty is not greater than that of the Debussy and Chopin pieces, al-

TABLE 1
Error Frequencies and Percentages

Piece	No. of Notes	Substitutions	Omissions	Intrusions	"Untied" notes
(a) Error frequencies, separately for the three performances of each piece[a]					
Schumann	4,570	2 2 4	30 38 37	23 22 30	56 46 43
Debussy	5,830	27 27 20	102 103 90	32 29 20	16 14 14
Chopin	15,180	34 32 28	262 214 200	156 135 154	[b]
Grieg	6,580	27 18 24	190 164 165	156 154 153	[b]
(b) Error percentages, for the three performances combined[c]					
Schumann	13,253	0.06	0.77	0.55	9.86
Debussy	16,907	0.42	1.69	0.46	2.07
Chopin	44,022	0.21	1.48	0.98	[b]
Grieg	19,082	0.35	2.63	2.35	[b]

[a]The error counts from P4's second performance are included in the totals for both the second and third performances, but are included only once in the error percentages.

[b]Untied notes were too infrequent to be listed separately; they are included among the intrusions.

[c]Intrusions are expressed as percentages of all notes played, untied notes as percentages of tied notes.

though it contains one particularly difficult measure.) The low error rates in the Schumann piece may reflect its even lower technical demands or its greater familiarity from listening. Interestingly, hardly any decrease was seen in the number of errors across the three performances of each piece, perhaps because they did not immediately follow each other. In separate analyses of variance on the error frequencies of each piece, the difference among performances did not approach significance.

INDIVIDUAL DIFFERENCES

Naturally, large individual differences among the pianists were apparent. Table 2 shows the individual total error frequencies and the ratios between intrusions and omissions (I/O ratio). The differences in accuracy were fairly consistent across the pieces and thus seemed to be related more to the pianists' level of skill and "quick study" ability than to whether they had studied a piece at some time in the past (indicated by boldface in Table 2). The average correlation among the individual error frequencies for the six pairs of pieces was 0.57, the highest correlation (.82, $p < .01$) being between Schumann and Debussy. The I/O ratios were less consistent; the average correlation was only .28, although there seemed to be some relation between Schumann and Debussy (.57, $p < .10$) and between Chopin and Grieg (.52). Note the high I/O ratios of P8 versus the very low ones of P4 and P5. It is not clear at this time whether the I/O ratio is interpretable in terms of some aspect of pianists' technical skill or personality (e.g., their relative willingness to take risks).

TABLE 2
Individual Differences in Total Error Frequencies and Intrusion/Omission (I/O) Ratios

Pianist	Schumann		Debussy		Chopin		Grieg	
	Total	I/O	Total	I/O	Total	I/O	Total	I/O
P1	29	0.00	29	0.25	58	1.35	33	0.65
P2	25	1.20	53	0.35	120	0.28	79	0.83
P3	97	2.39	106	0.38	240	0.72	190	1.30
P4	34	0.22	46	0.13	198	0.21	76	0.11
P5	9	0.20	16	0.08	75	0.44	77	0.49
P6	7	—	44	0.31	46	1.25	133	0.93
P7	22	1.00	23	0.08	65	0.35	105	0.63
P8	3	—	28	0.64	125	1.70	68	3.13
P9	48	0.22	56	0.07	137	0.64	89	1.41
P10	58	0.33	93	0.33	151	1.12	201	1.00

Error frequencies of pianists who had studied the piece at some time are in boldface.

ERROR LOCATION

Errors were not evenly distributed throughout each piece. Not surprisingly, they were concentrated in the technically more difficult parts, which contained more notes and were also more difficult to read. A high similarity in error patterns was noted for structurally similar parts of a piece. This is illustrated in Figure 1 for intrusions and omissions in the Chopin prelude. Total intrusion frequencies per bar are plotted as positive numbers, whereas omission frequencies are shown as negative numbers. The boundaries of major sections are marked by vertical lines. The vast majority of errors occurred in the middle section (bars 28–75). This section is divided into two parts (bars 28–59 and 60–75). The first part is again divided into two subparts (bars 28–43 and 44–59) which are virtually identical, and the second part has two subparts of considerable similarity (bars 60–67 and 68–75), except for the final bar. The similarity of the error patterns for these identical or similar passages is evident in the figure. The corresponding correlations are 0.78 (omissions) and 0.89 (intrusions) in the first part, and 0.70 (omissions) and 0.69 (intrusions) in the second part (with bars 67 and 75 excluded). On the other hand, the patterns of omission and intrusion errors showed little similarity. In fact, a negative relationship seems to exist, at least in bars 36–43 and the analogous bars 52–59: Omission er-

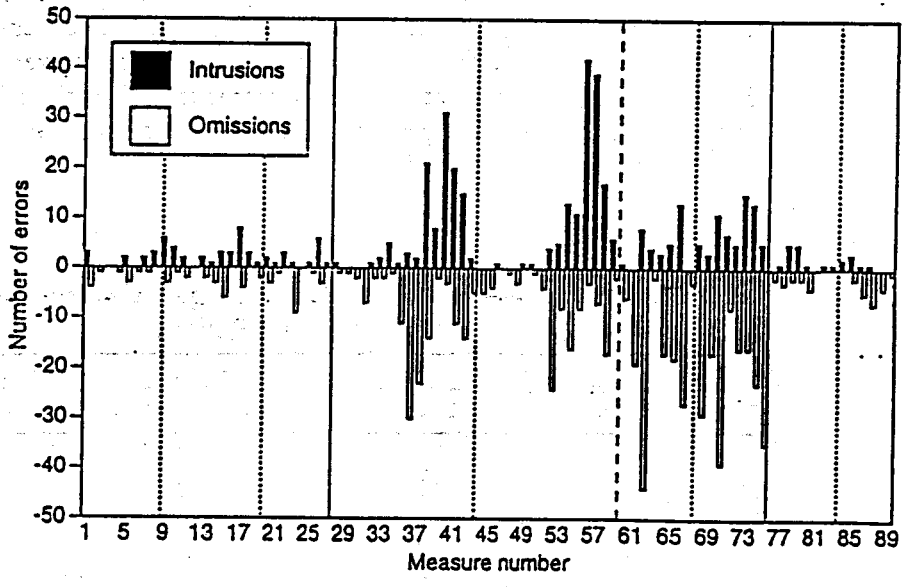


Fig. 1. Distribution of intrusions and omissions in Chopin's Prelude in D-flat Major. The vertical lines indicate major section boundaries.

rors first increase, then decrease as intrusion errors increase dramatically, and then increase again as intrusion errors decrease. This pattern parallels the dynamics of the music: The peak incidence of intrusion errors (bars 40 and 56) coincides with the peak of the crescendo that takes place during these bars.

A structural parallelism could also be seen in the Grieg piece, where the error patterns during the initial 8 bars and their subsequent repeat were extremely similar. Although the errors were not as strongly clustered as in the Chopin piece, a record number of omissions (113 total, a rate of more than 10%) occurred in bar 31, probably the technically most difficult bar in all the music used here. In the structurally more irregular Debussy piece, errors were concentrated in the last two thirds of the middle section (bars 19–27). In the Schumann piece, which has a very regular structure, errors were too sparse to reveal convincing structural correspondences. A more detailed description of the error patterns follows.

SCHUMANN: *TRÄUMEREI*

Substitutions

Two of the eight substitutions actually constituted an order reversal of two successive notes, the only such error in the present corpus and not a true substitution error. Five of the remaining six errors concerned the inner note B_2 in the left-hand part of the seven-note fermata chord at 22-2-1; four of these substitutions were D_3, which fits harmonically, although one was $B\flat_2$, a bad slip. The remaining error was a substitution of G_3 for $B\flat_3$ at 12-1-1, a harmonically appropriate but nevertheless fairly serious error because it alters the principal melody that occurs here in an inner voice.

Omissions

Twelve of the 101 omission errors[5] were due to apparent misinterpretations of the notation: At 8-3-1, where C_5 is preceded by a grace note of the same pitch, two pianists always played only a single note. (One effectively tied the two notes together, the other omitted the grace note.) These were the only omissions in the principal melody. Seven additional omissions, none of them repeated, occurred in secondary melodic voices that move in parallel with the principal (soprano) voice. Of the remaining 82 errors, 32 represented "incorrect ties"—failures to play a note that followed a note of the same pitch, played with the same finger. Most of these errors occurred

5. Note that the actual numbers of errors were smaller than the sums of the frequencies shown in Table 1(a), which include estimates of error frequencies for P4's missing third performance.

in the bass voice (C_3 at 3-2-1, 19-2-1, and 23-2-1; D_3 at 12-2-2), some in the tenor voice (C_4 at 7-4-2; C_4 at 23-2-1), all in the left hand. These errors tended to be repeated. This leaves 50 errors, all of which occurred in inner notes of chords. Eleven of them occurred with some consistency in the fermata chord (22-2-1). Nine were idiosyncratic to one pianist (F_3 in the chord at 1-2-1 and analogous positions). The remaining 30 represented various insignificant notes.

Intrusions

Forty-three of the 75 intrusions were due to a single pianist (P3). One common form of intrusion is due to accidentally "brushing" a key with the finger, which results in a tone of very low intensity (see also Moog & Rhea, 1990, p. 57). Nearly all of P3's intrusion errors were of that sort; some of them were temporally isolated and few were contextually appropriate. Surprisingly, however, several of them occurred repeatedly. Of the other pianists' 32 intrusions, many were also of relatively low intensity, and there was very little consistency, either within or between pianists. The only position in the music exhibiting intrusions by several pianists was again the fermata chord at 22-2-1. Only a few errors could be interpreted as anticipations or perseverations of contextual notes. The majority of the intrusions by pianists other than P3 provided reasonable harmonic fits; only 10 (31%) were judged by the author to be harmonically jarring.

Untied Notes

The large number of untied notes was unexpected. Even some pianists who otherwise played very accurately were inclined to ignore ties. The score contains 49 tied notes, but only 15 of them were subject to untying, often repeatedly but rarely with complete consistency: F_4 at 3-1-1 and 19-1-1, C_4 at 4-1-1 and 20-1-1, A_3 or A_4 at 7-1-1, C_4 at 7-3-1, D_4 at 11-1-1 and A_4 at 15-1-1, D_4 at 12-3-1 and A_4 at 16-3-1, F_4 at 23-1-1, D_4 at 24-1-1, and C_3 or C_4 at 24-3-1. Clearly, the pianists did not ignore all ties but only those in certain places. These were the ones where other notes played by the same hand, connected through legato to preceding notes, coincided with the tied note. There the hand could "lean on" other keys while the key of the tied note was restruck; elsewhere the hand would have had to be moved just to play the tied note. This observation suggests that "untying" was not due to misreadings but had technical causes. This is also consistent with the fact that the two ties that extended from the end of one system to the beginning of the next (15-1-1, 20-1-1) were not broken more often than their analogous uninterrupted occurrences (4-1-1, 11-1-1). Of course, all untied notes were contextually highly appropriate.

DEBUSSY: *LA FILLE AUX CHEVEUX DE LIN*

Substitutions

Of the 70 substitutions, 15 were caused by a misprint in the score: At 18-2-1, the first of the two ledger lines above the low note $B\flat_1$ was missing. This inadvertently created a small experiment on the reading of ambiguous notation: Would the pianists go by the vertical position of the note head and play the correct pitch, or would they go by the number of ledger lines and play $D\flat_2$ instead, a moderately jarring note? As it happened, five pianists did one thing and five the other, with complete consistency.

Another error committed by four pianists—consistently by two of them—was the substitution of $A\flat_2$ for $G\flat_2$ at 24-2-1, an inner note of a 5-note chord. The error was clearly context-induced because it changed the left-hand interval from a fourth to a fifth in a context of parallel fifths (which seems reasonably acceptable harmonically).[6] Other errors observed more than once were the substitution of $G\flat_3$ for $A\flat_3$ at 9-3-3, also an inner note of a chord and harmonically appropriate, and of F_3 for $F\flat_3$ at 15-1-1, an incongruent note resulting from neglect of an accidental in the score and thus apparently a misreading error, although it was not committed consistently by anyone. Most of the remaining errors were unique and occurred during the arpeggi and chords of bars 19–27, which represent the technically most difficult passages in the piece. None of the errors occurred in a principal melody or in the highest note of a chord. Of the 55 substitutions (discounting the 15 errors due to the misprint in the score), the author judged 20 (36%) to be harmonically jarring.

Omissions

Omissions were astonishingly frequent (a total of 286). The large majority occurred during bars 19–27, which constitute the second and third parts of the middle section, where also most substitutions were located. Repetition of errors across performances was common: 33 individual omissions were completely consistent, and 29 were partially consistent. Two kinds of omission may have been due to reading difficulties: The $A\flat_3$ at 19-3-1 and the analogous $A\flat_4$ at 20-3-1 occur inside dense clusters of notes and are somewhat difficult to make out in the score, but they are also difficult to play if the wrong fingering is chosen; they were consistently omitted by four pianists. A single omission of an isolated melody note represented an "incorrect tie"—a failure to play a repeated pitch. There was one unique

6. A contextually more appropriate misreading would have been to substitute C_1 for $D\flat_2$ instead, based on parallelism with the voices in the right hand (cf. 25-2-1). No pianist made that substitution, however, probably because of the prevalence of black key notes in the immediate context.

instance of a completely omitted four-note chord. Five other omissions, all unique, were judged by the author to be conspicuous; all others seemed relatively insignificant, being inner notes of right-hand chords.

Intrusions

Intrusions, too, were most prevalent in bars 19–27. Most were unique and occurred inside chords, often providing a harmonic completion. The only error committed repeatedly by several pianists was the intrusion of $G\flat_3$ at 23-2-1, which is a good example of this chord-filling tendency. Only 29 (36%) of the 80 intrusions were judged by the author to be harmonically inappropriate. (One rare fumble involved five intrusions in close succession.) As in the Schumann, P3 contributed the largest number of intrusions, but here only a few were of the "brushed key" variety.

Untied Notes

Untied notes were not abundant (a total of 41), even though the score contains many tied notes. Most of these, however, are conspicuous chords that are not likely to be struck twice. In fact, the untying was mainly restricted to three positions: $B\flat_2$ at 20-1-1 (a tie across a page break), $B\flat_3$ at 20-3-1, and $D\flat_3$ at 24-1-1. Curiously, $B\flat_3$ at 19-3-1 and $B\flat_4$ at 20-3-1, although quite analogous to $B\flat_3$ at 20-3-1, were never untied. There were seven instances of complete consistency and four of partial consistency. One pianist consistently played the two right-hand notes in the final chord at 39-1-1, which have an unusual "open-ended" tie in the score but clearly are not intended to be sounded again.

CHOPIN: PRELUDE IN D-FLAT MAJOR

Substitutions

The large majority of the 94 substitutions occurred in the middle section, especially bars 61–71, where full chords and complex harmonies predominate and the characteristic pulsing eighth-notes appear in changing inner voices. Two pianists were particularly prone to misplay some of these notes by repeating the lowest note of the preceding right-hand chord instead; thus they consistently played $D\sharp_4$ instead of $G\sharp_4$ at 61-4-2, 62-4-2, and 69-4-2, and they often played E_4 instead of $G\sharp_4$ at 62-1-2, 62-2-2, and 70-2-2. A number of similar but less consistent substitutions were done by other pianists. Although these errors can be regarded as perseverations or anticipations of neighboring pitches (Palmer & van de Sande, 1993), they are more readily explained as technical simplifications: It is easier to play repeated notes with the thumb than with the index finger while holding

down a chord with the other fingers. Of course, all these substitutions fit the harmonic context perfectly. Only a single substitution occurred in the soprano melody; five were in melodic middle or bass voices. Only 15 of the substitutions (16%) were judged to be contextually inappropriate.

Omissions

Nearly all of the 676 omissions occurred in the middle section (see Figure 1), particularly where right-hand chords with repeating inner notes were involved (bars 36–42 and 52–75). The vast majority of the omissions concerned these repeated eighth notes. Most of the omissions occurred when a note an octave higher or lower was sounded simultaneously. For example, a particularly common omission (made consistently by several pianists) was $G\sharp_4$ at 62-1-1 and 70-1-1, which facilitates the passage technically a great deal. Sometimes, however, repeated notes occurring in isolation were omitted, although never consistently. On the basis of such score-based observations (not actual listening), the author judged 62 omissions (9%) likely to be conspicuous; 50 of them were isolated repeated eighth notes. Nearly all of these omissions were unique. Only three occurred in a melody voice (two in the soprano and one in the bass).

Intrusions

The 445 intrusions included 17 untied notes, always $C\sharp_4$ at 38-1-1 and/or 54-1-1, two of the very few tied notes in the score. As was noted in connection with Figure 1, intrusions were especially frequent in the fortissimo passages of the middle section. A striking difference was also seen in their relative contextual appropriateness at different dynamic levels. Bars 1–37, 43–53, 59–70, and 76–89 are the quieter sections of the music, whereas bars 38–42, 54–58, and 71–75 are the fortissimo and forte sections. The soft sections (74 bars total) contained 171 intrusions, of which 58 (34%) were judged to be harmonic misfits. The loud sections (only 15 bars total) contained 257 intrusions, of which 217 (84%) were judged to be contextually inappropriate. Whereas the errors in the soft sections consisted mostly of added inner notes in chords and added pulsed eighth notes where they were absent in the score, the errors in the loud sections consisted predominantly of stray notes inside the powerful fortissimo chords, evidently caused by accidentally touching keys with the free fingers. Most of these added tones had very low intensities relative to the very loud tones that surrounded them; thus they probably just added a blur or fuzziness to the musical texture (a phenomenon commented on by Moog & Rhea, 1990). Despite the great frequency of intrusions, only a few were made consis-

tently by one pianist or were shared by several pianists (except for the two untied notes).

GRIEG: *EROTIK*

Substitutions

Nearly half of the 68 substitutions in this unfamiliar piece were due to one pianist (P3), who admitted being a poor sightreader. Most substitutions were idiosyncratic, and few were committed more than once by the same pianist. In fact, the only error exhibiting complete consistency was P3's substitution of $B\flat_4$-$B\flat_5$ for A_4-A_5 at 7-2-4 and 15-2-4, which is a melodic error but fits reasonably well into the melodic-harmonic context. Only 17 (2.5%) of the 68 errors were judged to be jarring, most of them only moderately so. Only one of them occurred in the principal melody.

Omissions

Omissions were abundant. Pianists exhibited considerable consistency across their three performances, and the same omission errors were often committed by several pianists. They were predominantly inner notes of chords. Examples of frequently occurring errors are F_4 in 3-2-3 and 11-2-3; F_3 in 7-1-1 and 15-1-1; D_5 in 7-2-1 and 15-2-1; D_3, C_4, F_4, or G_4 in 8-1-1 and 16-1-1; $B\flat_3$ in 20-1-1; and A_3, C_4, F_4, or A_4 in 29-1-1-1. In bars 21–28, the lower notes of the syncopated chords tended to be dropped: G_3 and/or $B\flat_3$ in 21-1-3; $B\flat_3$ and/or D_4 in 23-1-3; $D\flat_4$ and/or F_4 in 25-1-3; $C\sharp_4$ and/or G_4 in 26-1-3 in 27-1-3; F_4 and/or G_4 in 28-1-3. In bars 29–31, omission of the lower notes of the two-note syncopated accompanying chords was the most common error, especially C_4 in 31-2-1-2, F_3 in 31-2-2-2, and C_4 in 31-2-3-2. Thus the omissions resulted in a slight thinning of Grieg's thick chordal textures, which was unlikely to be perceptually salient. Only 7 of the 496 omissions occurred in the soprano voice and were correspondingly conspicuous.

Intrusions

The high frequency of intrusions was perhaps the clearest indicator of the pianists' discomfort with the unfamiliar piece. As in the other pieces, intrusions were more idiosyncratic and less consistent than omissions. Seven of the errors represented untied notes, for which there were only two opportunities in this music (G_2 at 21-1-1; B_3 at 28-2-3). The most common error was a restriking of the syncopated accompaniment together with the melody notes in bars 21–31, an error that of course fit the harmonic con-

text. A substantial number of the 461 intrusions (186 or 40%), however, represented "slips" that did not fit the context well. Only seven of these errors were completely consistent, compared with 30 among the contextually appropriate intrusions.

SUMMARY

The data support the four main hypotheses stated in the Introduction: (1) All types of errors occurred almost exclusively in subsidiary voices, especially in the inner voices of chords. (2) The errors occurred mainly in positions where many notes coincided. (3) The majority of the errors were judged to be contextually appropriate, with the exception of intrusions in very loud passages. (4) Omissions were more consistent than intrusions, except for untied notes, which were often consistent. Consistent errors were nearly always contextually appropriate or, in the case of omissions, likely to be inconspicuous. Contextually inappropriate errors were usually inconsistent. In addition, (5) intrusions tended to be of relatively low intensity.

Error-Detection Experiment

The purpose of the error-detection experiment was to confirm the author's general impression that most errors were difficult to hear and to provide evidence bearing on the hypotheses stated in the Introduction. (Hypothesis 4 was less relevant here.)

METHOD

Ten complete performances of the Chopin prelude—each pianist's last performance—were reproduced from the original MIDI files on a Roland RD-250s digital piano and recorded onto digital tape.[7] Eight undergraduate students, all skilled pianists, served as paid listeners. Two had recently studied Chopin's 24 preludes, and one had performed them in recital; the other six were familiar with the piece from listening only. Each pianist listened to the 10 performances in a random order over Sennheiser HD 540II earphones at a comfortable intensity, with a break after the fifth performance. Ten copies of the score were provided. The task was to mark all pitch errors in the score by circling any wrong or missing note(s) and by making an open circle where an added note was heard, without regard to its pitch. It was emphasized that it was less important to circle the correct note(s) than to indicate that an error had been heard and that, in the case of uncertainty, it was all right to make a large circle around several notes. Errors that did not concern pitch (e.g., rhythmic errors, poor articulation or pedaling) were to be ignored.

7. For various reasons, one being the reliable response of the digital piano to low MIDI velocities, this form of reproduction was preferred over an acoustic recording from the original Disklavier. It seemed unlikely that the relative detectability of errors would depend on the instrument used. The only change made in the original MIDI files was the removal of the soft-pedal instructions because they created unnaturally large dynamic differences on the digital piano.

TABLE 3
Summary of Error Detection Scores

Type of Error	Total	Detected	Percentage
Substitutions	28	19	68
Omissions	198	81	41
Intrusions	154	38	25
Total	380	143	38

A liberal criterion was adopted in scoring the responses. A substitution or omission was considered detected when any note in the same chord or in an adjacent position had been circled, or when a wide circle included the correct note. An intrusion was considered detected when an open circle was in the vicinity of the correct position or when a note close in pitch had been circled. When a note far in pitch from the intrusion had been circled, it was scored as a false alarm (omission) response. When two or more errors of the same kind occurred in the same position, all were considered detected if one was marked.

RESULTS

Table 3 shows how many errors were detected by at least one listener. It can be seen that only 38% of all errors were detected; 62% were never reported by any listener. Substitutions were detected more often than omissions, which in turn were detected more often than intrusions. Only 6 of 380 errors were reported by all 8 listeners, 3 errors by 7 listeners, and 5 errors by 6 listeners. Individual detection scores ranged from 25 (7%) to 83 (22%). The two pianists who had recently studied the piece had the highest scores. The number of false-alarm responses per listener ranged from 6 to 49.

Substitutions

Only 4 of the 28 substitutions had been classified as contextually inappropriate by the author. Three of these were among the most frequently detected errors: the double error of $B\flat_4$ and $D\flat_4$ instead of $C\sharp_4$ and E_4 at 66-1-1 (eight reports), and E_4 instead of $A\flat_3$ at 64-4-1 (six reports).[8] The fourth error judged to be potentially jarring ($F\sharp_2$ instead of $A\flat_2$ at 57-4-2) was never reported, probably because of its low pitch, short duration, and coincidence with other intense tones. Examples of other substitutions never detected are $D\flat_4$ for $G\sharp_4$ at 60-1-1; $G\sharp_4$ for $C\sharp_4$ at 71-2-1 and 71-2-2.

Omissions

Twenty-four of the 198 omissions had been classified by the author as likely to be noticed; all of them were in fact reported at least once. Four of

8. Technically, the second error, too, was a double substitution ($C\sharp_4$ and E_4 instead of $A\flat_3$ and $C\sharp_4$), but it was counted as a single error because only one pitch changed.

them were heard by all eight listeners: $A\flat_3$ at 4-4-2, $A\flat_2$ at 9-1-1, $D\flat_3$ at 80-1-1, and $B\flat_4$ at 11-2-1; the first three are prominent bass notes, the last a melody note. A double omission in position 9-2-1 ($E\flat_4$ and C_4) was reported by seven listeners. Representative examples of omissions never detected are B_3 at 58-4-2 (the lower note in an octave); $G\sharp_4$ at 37-3-2 (the upper note in an octave); $G\sharp_4$ at 69-4-1 (a note inside a full chord).

Intrusions

Frequently reported intrusions were $E\flat_5$ at 17-1-2, a jarring note in the melody register (eight reports); a whole four-note chord at 34-4-2 (seven reports); the double intrusion $E\flat_4$ and $C\flat_3$ at 13-3-1, two harmonically appropriate notes (seven reports); F_4 at 38-3-1, a dissonant note in a chord (seven reports). The author had judged 96 of the intrusions to be harmonically jarring; only 18 (19%) of these were detected. Of the 58 contextually appropriate intrusions, however, 20 (34%) were detected. This was contrary to expectations (Hypothesis 3), but it can be explained by the different dynamic contexts in which these errors occurred. Contextually inappropriate intrusions occurred mostly during loud passages, where they tended to be masked by other loud tones. A modest negative correlation of $-.43$ ($p < .001$) between detectability and the difference between the MIDI velocities of the intrusion error and the simultaneous highest-pitch note, as well as a similar correlation with the absolute MIDI velocity of the latter, but no correlation with the MIDI velocity of the intrusion itself. Even very soft intrusions were sometimes detected: Of 49 intrusions with MIDI velocities below 10, ten (20%) were reported. One striking finding was that intrusions were much more detectable in some performances than in others: Ten of P5's 11 intrusions were reported, most of them more than once, whereas none of P8's 31 intrusions was ever detected. The main difference was that P5's intrusions occurred mostly in the first part of the piece whereas P8's occurred mainly in the intense passages of the middle section.

False Alarms

False-alarm responses, 137 in all, were distributed over 105 positions. Eleven of them were "open circle" responses, indicating intrusions; all the others indicated omissions (or substitutions). Not surprisingly, the cause was often a very soft note, especially when it coincided with a sustained chord; the repeated notes after the subito piano in bars 43 and 59 are examples. A few false alarms were triggered by timing or pedaling errors.

Summary

The results confirm that most errors in these performances were difficult to hear. Only a handful of errors was detected by most listeners; these in-

cluded the very few errors that occurred in the melody or bass voices (with one exception noted above). The author's score-based predictions of the detectability of omissions were borne out, except that the detectability of intrusions depended more on their dynamic context than on their judged contextual appropriateness. Masking by other tones seems to be the primary cause of why intrusions are difficult to hear.

Conclusions

These results may be regarded from two perspectives. From the pianist's perspective, it is good to know that only a fraction of the inaccuracies committed are ever noticed by an audience. From a listener's perspective, it is interesting (although perhaps disquieting) to learn that performances often are much less accurate than they seem to the ear.

Performers, who monitor their own movements as well as the resulting sounds, may notice errors that listeners do not hear. MIDI registration in addition detects errors that performers may not notice. Thus errors can be defined and counted at different levels in the process of musical communication. Arguably, however, the most important level is the listener's perception. Errors that are not noticed are insignificant for all practical purposes. The present study suggests that most objectively registered errors are not detected and hence are inconsequential.

This conclusion may well hold regardless of the specific performance situation. The errors discussed here were committed by skilled young pianists in a "quick study" situation. Certainly the absolute error frequencies would have been much lower if the pianists had been experienced concert artists or if the pieces had been rehearsed extensively, whereas error frequencies would have been higher if the pianists had been high school students or if the pieces had been sightread for the first time. However, there is no obvious reason why the proportion of perceptually salient errors should change dramatically. If anything, it might decrease as the number of errors decreases: As performance becomes more sophisticated, the errors (if any) probably become more sophisticated, too.

The present results demonstrate that performance errors vary along a continuum of perceptual salience. Errors are categorical only at the level of MIDI events in relation to the printed score, but at the level of the listener's perception they cease to be categorical. The listener's musical experience, knowledge of the music, availability of the score, level of attention, and other factors determine a perceptual criterion or threshold that admits only a certain proportion of errors to consciousness. Even errors that have been detected still vary in degree of severity, and a listener probably could rate them accordingly. Errors then are a matter of degree. In that sense, pitch errors are no different from "errors" of timing, tempo, dynamics, or inter-

pretation more generally. Not long ago, it was common for pianists to double octaves, fill chords, and take other liberties that the contemporary ethos of "faithfulness to the score" has banned. Some kinds of perceptually insignificant pitch errors, such as the observed tendencies to fill chords or undo ties, could be viewed as spontaneous "liberties," some of which may even enhance the aesthetic effect of the music. Whether they are considered errors at all depends on whether the object of faithfulness is taken to be the printed score or the music as heard.[9]

References

Huron, D. (1989). Voice denumerability in polyphonic music of homogeneous timbres. *Music Perception, 6,* 361-382.

Moog, R. A., & Rhea, T. L. (1990). Evolution of the keyboard interface: The Bösendorfer 290 SE recording piano and the Moog multiply-touch-sensitive keyboards. *Computer Music Journal, 14,* 52–60.

Palmer, C., & Holleran, S. (1994). Harmonic, melodic, and frequency height influences in the perception of multivoiced music. *Perception & Psychophysics, 56,* 301-312.

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

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

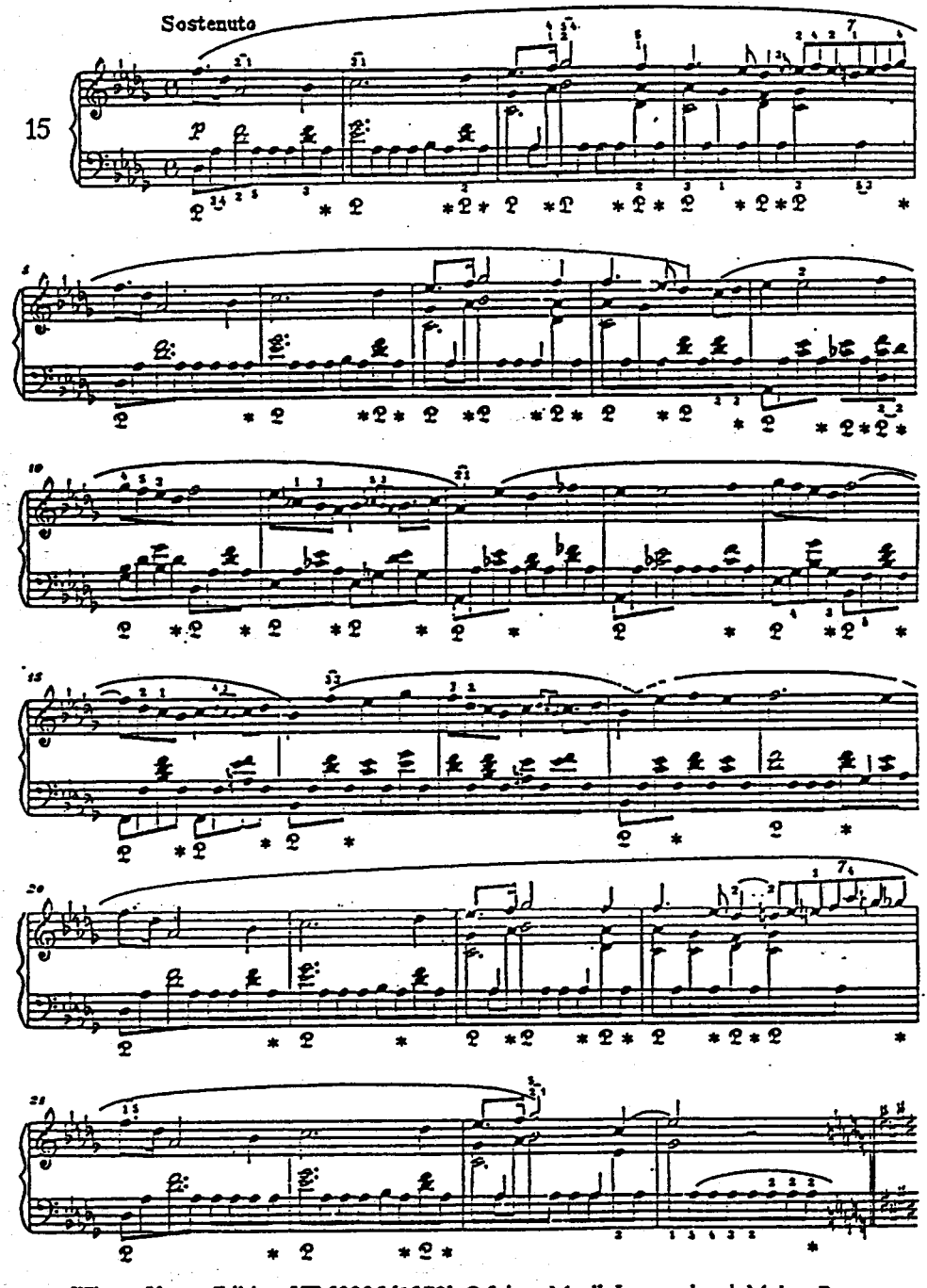
Rasch, R. (1978). The perception of simultaneous notes such as in polyphonic music. *Acustica, 40,* 21–33.

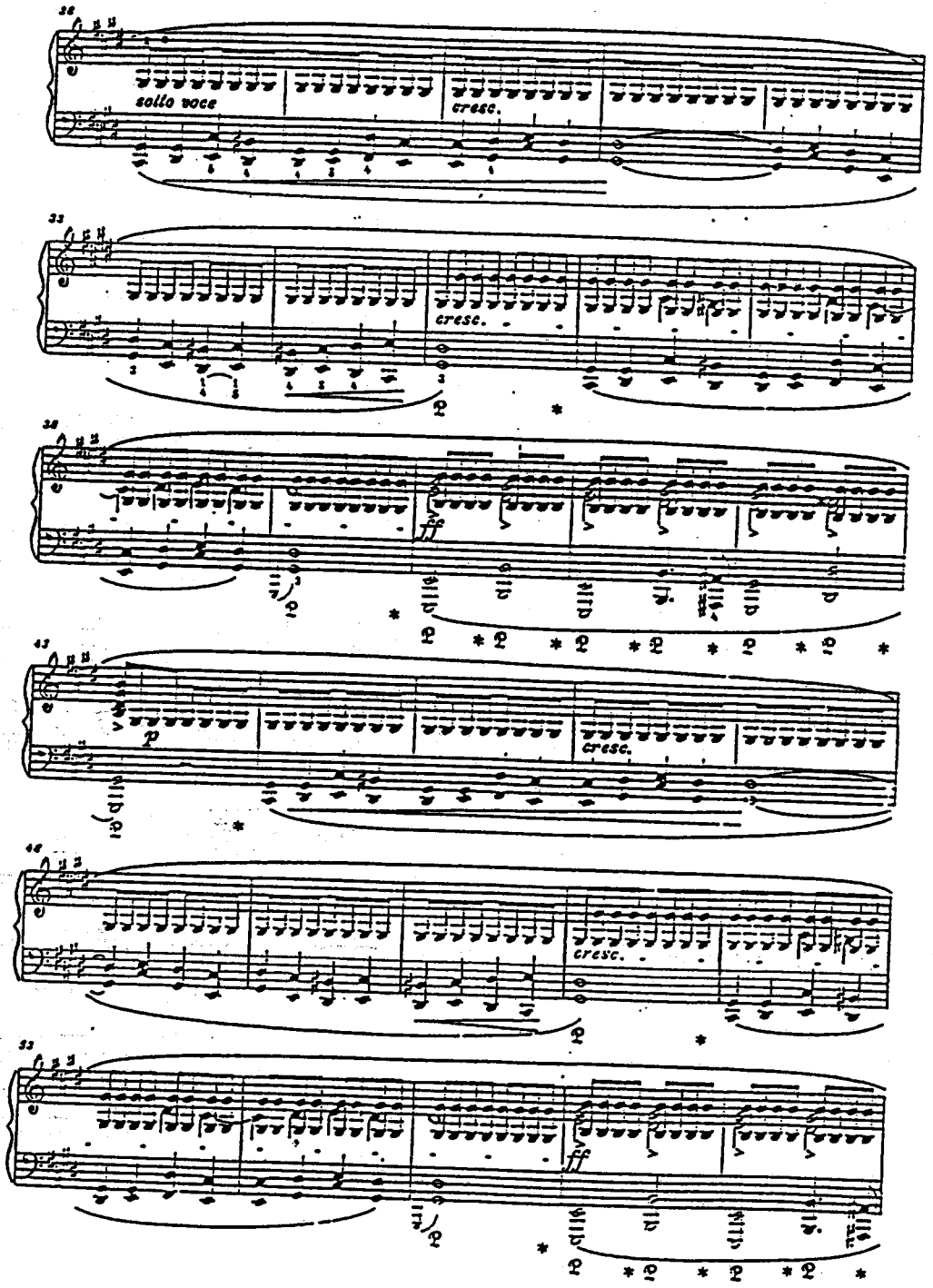
Sloboda, J. A. (1985). *The musical mind: The cognitive psychology of music.* Oxford: Clarendon Press.

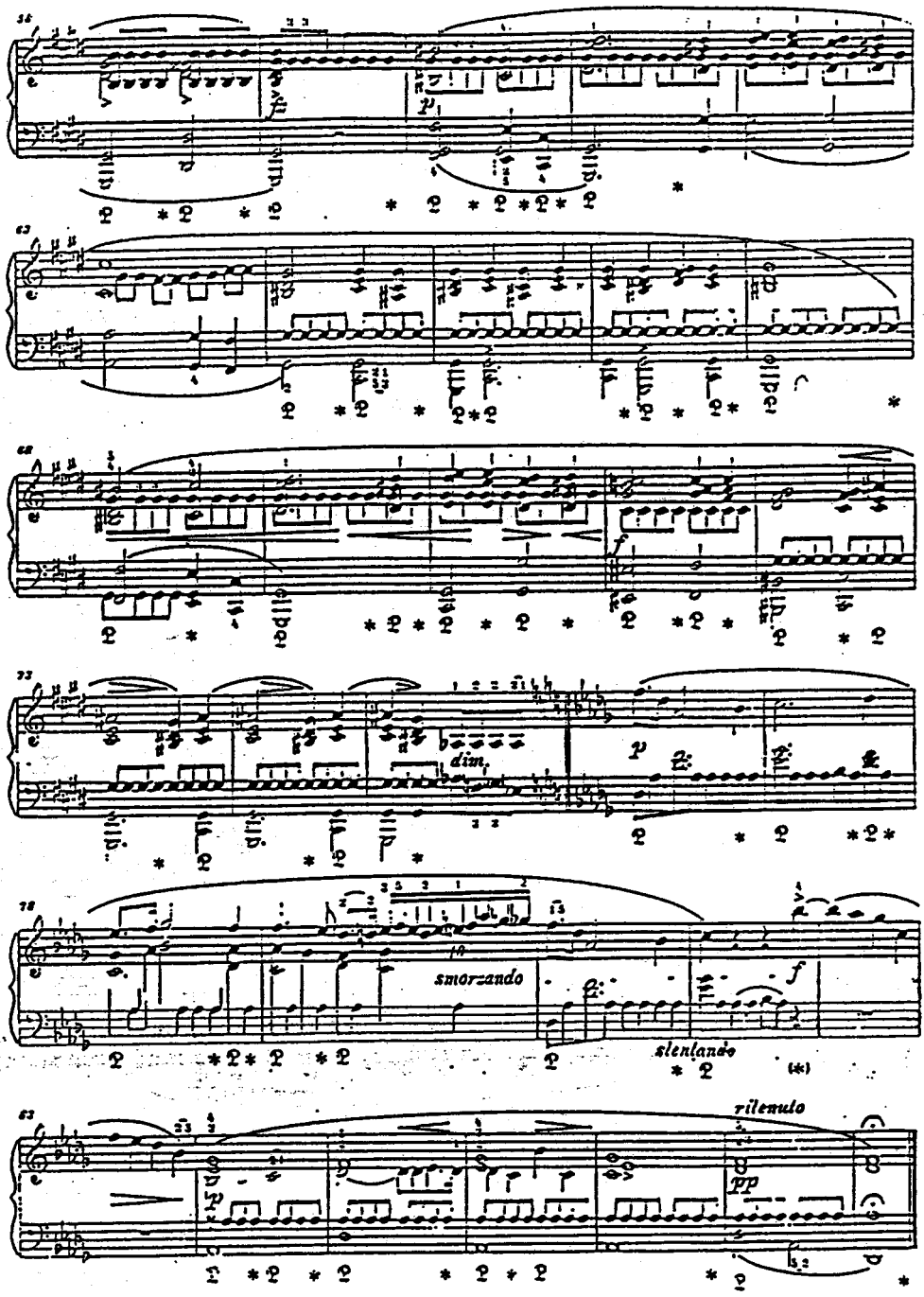
Taruskin, R. (1995). *Text and act: Essays on music and performance.* New York: Oxford University Press.

9. This research was supported by National Institutes of Health grant MH-51230. I am grateful to Charles Nichols, Linda Popovic, Ilan Berman, and Mary Wilkins for assistance, to Jonathan Berger (Director, Center for Research in Music Technology, Yale University) for permission to use the Yamaha Disklavier, to Caroline Palmer, Nigel Nettheim, Linda Popovic, and two anonymous reviewers for their comments on an earlier draft of the manuscript, and to the pianists for lending their skills to this project.

Appendix
Frédéric Chopin: Prelude in D-flat Major, op. 28, no. 15




sotto voce
cresc.
cresc.
ff
p
cresc.
cresc.
ff


dim.
smorzando
slentando
rilentato