Perception & Psychophysics
2006, 68 (3), 362-376

Effects of hearing the past, present, or future during music performance

PETER Q. PFORDRESHER
University of Texas, San Antonio, Texas

and

CAROLINE PALMER
McGill University, Montreal, Quebec, Canada

Three experiments were performed to explore the effects of mismatches between actions (keypresses) and the contents of auditory feedback (pitch events) during music performance. Pianists performed melodies from memory during altered auditory feedback that was synchronized with keypresses but matched the pitch of other sequence events. Feedback direction was manipulated by presenting pitches that matched events intended for the past (delays; Experiments 1 and 3) or the future (prelays; Experiments 2 and 3). Feedback distance was manipulated by varying the absolute separation between the current event and the location of the feedback pitch. All alterations disrupted the accuracy of performance (pitch errors) more so than timing. Serial-ordering errors indicated confusions among proximal and metrically similar events, consistent with the predictions of an incremental planning model (Palmer & Pfordresher, 2003). Patterns of serial-ordering errors suggested that performers compensate for the disruptive effects of altered feedback by changing event activations during planning.

The production of auditory sequences, such as music and speech, simultaneously involves action and perception. A common technique for understanding perception–action relationships is to examine the influence of auditory feedback on production.[1] Many, although not all, feedback alterations disrupt production (for reviews, see Finney, 1999; Howell, 2004a, 2004b; Yates, 1963). We propose that disruption from altered feedback occurs when the feedback matches events being planned for production at other sequence positions. We conceptualize planning of sequential behavior as the use of a memory representation to prepare events for production. Representations of sequential behaviors are typically considered to be hierarchical (R. Cooper & Shallice, 2000; Dell, 1986; Lashley, 1951; MacKay, 1987; Miller, Galanter, & Pribram, 1960; Rosenbaum, Kenny, & Derr, 1983). One implication of hierarchical representations is that both past and future events can be simultaneously accessible if they are linked at a higher level. To address the link between planning and feedback, therefore, one should examine the influence of auditory feedback that repeats past events or anticipates future events. Unfortunately, in past research only the influence of past events has been examined.[2]

We will report three experiments in which pianists performed short melodies while the pitch contents of auditory feedback (i.e., the pitches sounded at each keypress) were altered to match pitches intended for nearby sequence positions. We will describe here a new methodology that allowed us to manipulate the direction and distance of feedback alterations. *Feedback direction* refers to whether altered auditory feedback pitches correspond to past events (delays) or future events (prelays). *Feedback distance* refers to the absolute separation (past or future) between the current position and the planned position of auditory feedback (one to three events, in the altered feedback conditions). Figure 1 shows the influence on a single produced event of each kind of feedback manipulation used in the experiments reported here.

The Role of Auditory Feedback

What is the link between the planning of actions and the effects of auditory feedback? Early accounts suggested that altered feedback disrupts feedback control of sequence production (see, e.g., Black, 1951; Chase, 1965; Fairbanks, 1954; Fairbanks & Guttman, 1958; Lee, 1950). However, such views are unlikely, given the time course of auditory feedback (Borden, 1979), as well as evidence that some alterations fail to disrupt performance (e.g., Finney, 1997; Pfordresher, 2005) and that performance often pro-

This research was sponsored in part by NSF Grant BCS-0344892, awarded to P.Q.P., and by the Canada Research Chairs program, NSERC Grant 298173, and NIMH Grant R01-45764, awarded to C.P. Portions of these results were presented at the 2001 Annual Meeting of the Psychonomic Society and the 2002 Auditory Perception, Cognition, and Action meeting. We thank Grant Baldwin for conducting experiments and Steven Finney for generous technical assistance. We also thank Bruno Repp and two anonymous reviewers for their thoughtful comments on an earlier version of the manuscript. Correspondence concerning this article should be addressed to P. Q. Pfordresher, Department of Psychology, Park Hall, University of Buffalo, Buffalo, NY 14260, or to C. Palmer, Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montreal, PQ, H3A 1B1 Canada (e-mail: peter.pfordresher@gmail.com or caroline.palmer@mcgill.ca).

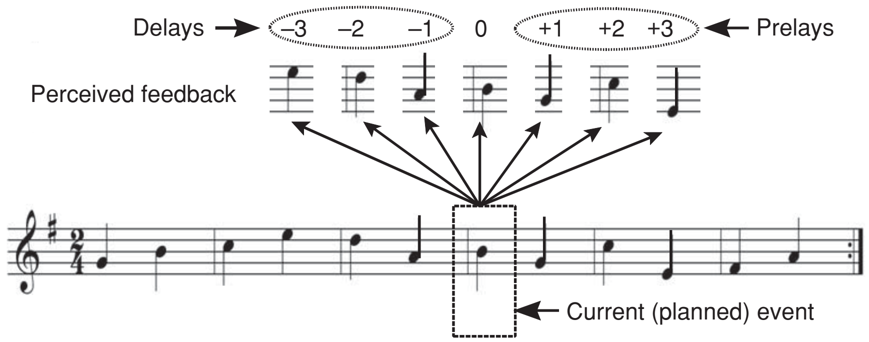

Figure 1. Depiction of auditory feedback manipulations created during music performance. Music notation for one of the stimulus melodies is shown at the bottom. Arrows above the music notation indicate which feedback pitch was heard upon production of the current event in all the possible feedback conditions.

ceeds unhindered when feedback is absent (Finney, 1997; Finney & Palmer, 2003; Repp, 1999).

Another view holds that alterations of auditory feedback do not disrupt planning but, instead, influence execution by virtue of asynchronies between movements and the amplitude contour of the sound, possibly because the delayed sound creates a separate, interfering signal (e.g., Howell, 2004a, 2004b; Howell & Archer, 1984; Howell & Au-Yeung, 2002; Howell, Powell, & Khan, 1983). This view is similar to the rationale behind the motor program hypothesis, in that production does not rely on feedback for error correction (for reviews, see Rosenbaum, 1991; Schmidt & Lee, 1999). Research has shown disruption of production when auditory feedback is asynchronous with production (first documented by Black, 1951; Lee, 1950), but not when the contents of feedback (e.g., musical pitches) are altered. For instance, pianists were not disrupted when feedback pitches that sounded in synchrony with keypresses (i.e., feedback contents) were randomized (Finney, 1997), and asynchronous speech feedback yielded similar performance when the contents (phonemes) were changed to square wave tones (Howell & Archer, 1984).

More recent evidence, however, has documented disruption of production when feedback contents are altered and the timing is unaltered (i.e., is synchronized). Pfordresher (2003) tested whether alterations of feedback contents (i.e., synchronous feedback) would disrupt production when altered feedback events were related to past produced events. Piano performances were disrupted when altered feedback contents matched the pitches of events produced one to three keypresses in the past, called *serial shifts* (see Houde & Jordan, 1998, and Müller, Aschersleben, Esser, & Müsseler, 2000, for similar manipulations of speech feedback).[3] Serial shifts primarily increase error rates, in contrast to alterations of auditory feedback synchrony, which primarily influence the timing of actions (Pfordresher, 2003).

We propose that altered auditory feedback disrupts production because the perception of auditory feedback and the planning of actions share a common memory representation of the sequence (Hommel, Müsseler, Aschersleben, & Prinz, 2001; MacKay, 1987; cf. Kalinowski & Saltuklaroglu, 2003). This proposal is similar to ideomotor theory (Greenwald, 1970; James, 1890), in that it suggests that actions are linked to planned outcomes. Planning (the mental activation of sequence events during retrieval) is assumed to be incremental: At any moment during production, some (but not all) events in the sequence are active or mentally accessible (Kempen & Hoenkamp, 1987; Palmer & Pfordresher, 2003; M. Smith & Wheeldon, 1999). Disruption occurs when feedback adds activation to events intended for noncurrent serial positions. Thus, auditory feedback events that are serially shifted disrupt production because feedback matches events that are active but are not intended for the present location (e.g., Pfordresher, 2003), whereas auditory feedback that presents randomized pitches fails to disrupt because feedback events match none of the accessible events (e.g., Finney, 1997).

Planning and Event Accessibility

Event accessibility in sequence production is often examined in analyses of serial-ordering errors: errors in which an event intended for elsewhere in the sequence is produced. The majority of serial-ordering errors (also called *movement* errors) originate from other sequence events (e.g., Fromkin, 1971; Garrett, 1980) and are often caused by confusions among similar sequence events (Conrad, 1965; Healy, 1974). Serial-ordering errors in sequence production (including speech and music) reveal both *directional* and *distance*-based constraints on the accessibility of events during planning (see Palmer & van de Sande, 1993, 1995, for further discussion).

The anticipatory proportion of errors (AP = number of anticipatory errors/number of anticipatory and perseveratory errors) measures the directional characteristics of planning in sequence production (Dell, Burger, & Svec, 1997; Vousden, Brown, & Harley, 2000). AP typically reveals an anticipatory bias: Errors tend to anticipate future events more often than they repeat past events (Dell et al., 1997). The anticipatory bias may be diagnostic of fluency in normal production; AP typically increases as error rates decrease (termed the *general anticipatory effect*; Dell et al., 1997). In production with normal feedback, the anticipatory bias may arise from postoutput suppression, as implemented in a recent model of sentence production (Vousden et al., 2000; cf. MacKay, 1987).

Serial-ordering errors also reveal distance-related constraints on planning. An event may be more or less accessible on the basis of its separation (in number of events) from the current event (e.g., Garcia-Albea, del Viso, & Igoa, 1989; Palmer & Pfordresher, 2003; Palmer & van de Sande, 1993, 1995). Movement gradients, or error proportions plotted by distance, measure event accessibility as a function of absolute distance (past or future) from the current event (Palmer & Pfordresher, 2003; Vousden et al., 2000).

The range model (Palmer & Pfordresher, 2003) predicts distance-related characteristics of planning. Event accessibility in the range model spans from the current event to nearby sequence events in the past and future, forming a gradient of event activations across the sequence. Activations are determined, in part, from similarity relationships among sequence events based on metrical accent strengths. In time-based sequences, such as music and speech, metrical accents guide production and perception by forming a hierarchical frame of alternating strong and weak beats (e.g., G. Cooper & Meyer, 1960; Lerdahl & Jackendoff, 1983; Liberman & Prince, 1977; Palmer & Krumhansl, 1990). Figure 2 shows metrical accents in grid notation below one of the stimulus melodies used in the present experiments; *X*s beneath musical events indicate the accent strength associated with each position. The range model predicts that events that share accent strengths with the current event are more active in planning, on the basis of their similarity, than are events associated with different accent strengths. Thus, an error at Position 7 (highlighted) would likely arise from Position 5 or 9, both of which are nearby and share the same accent strength (strong accent).

The range model formalizes metrical similarity between the current sequence event (position *i*) and other events separated from it by distance *x* in the following way:

$$\mathbf{M}_{x,i} = 1 - \frac{|\mathbf{m}_i - \mathbf{m}_{i+x}|}{\mathbf{m}_i + \mathbf{m}_{i+x}}, \tag{1}$$

where $\mathbf{M}$ is an array of similarity relationships between each current event (i) and surrounding sequence positions at different distances (x) and $\mathbf{m}$ is a vector of metrical accent strengths for individual events. Similarity relationships are straightforward in a simple binary metrical framework such as that shown in Figure 2: Each event is similar to events at distances that are multiples of 2 but is dissimilar to other events. Similarity relationships among pairs of events are averaged across sequence positions (i), to generate model predictions as a function of distance (x).

Weights can be assigned in the range model to different levels of the metrical hierarchy. Metrical weights reflect the salience of different hierarchical levels in meter—for

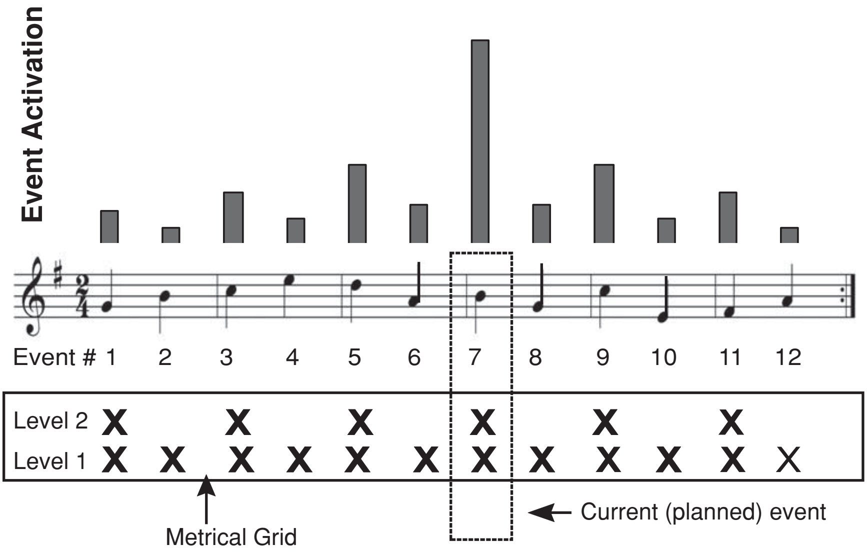

Figure 2. Range model predictions of event accessibility during performance (Palmer & Pfordresher, 2003). Top: Bars above the music notation are hypothetical predictions of the range model when Event 7 (rectangle) is the currently produced event. Center: Music notation for one of the stimulus melodies. Bottom: The metrical grid indicating metrical accent strength. The number of *X*s that align under each event indicates the number of levels in the metrical hierarchy that are accented at that position (one *X* = Level 1 only = weak accent; two *X*s = Levels 1 and 2 = strong accent).

instance, the metrical level at which people clap or tap their foot to music (see Duke, 1989; Parncutt, 1994). Metrical weights in the model are adjusted in the following way:

$$m_i = \sum_{j=1}^{k} w_j \cdot g_{j,i},$$

$$\text{where} \sum_{j=1}^{k} w_j = 1 \text{ and } 0 < w_j < 1. \qquad (2)$$

The w parameter gives a weight to each metrical level j (represented by an X in Figure 1). The g parameter indexes whether a particular metrical level is represented ($g = 1$) or not ($g = 0$) at a given sequence position, and m_i is the resulting metrical accent at a given position. For the current event highlighted in Figure 1, both metrical levels are present ($g_{1,7} = 1$ and $g_{2,7} = 1$). The default weight for each metrical level is $1/k$ (1/number of metrical levels), which designates equal weights for each metrical level ($w_1 = w_2 = .5$ for the grid in Figure 1). In most Western tonal music, $k = 2$ to 4 levels (see Lerdahl & Jackendoff, 1983; Palmer & Krumhansl, 1990; Palmer & Pfordresher, 2003).

The range model combines this similarity metric with a serial proximity component, which determines event accessibility on the basis of each event's separation from the current event. The proximity component depends on the produced tempo of the sequence (t, a fixed parameter) as well as on an estimate of working memory capacity (a, a free parameter). The product of these two components predicts the accessibility of events in a planning increment; one example is shown above the notation in Figure 2. The probability of an error's arising from any sequence distance diminishes for positions at greater serial separations from the current event (serial component), and errors should be relatively more likely to arise from distances that are separated from the current event by multiples of 2 (metrical component). Furthermore, if event distinctiveness from metrical accents increases, differences in error probability between distances that are multiples of 2 and other distances should grow larger (see Palmer & Pfordresher, 2003, for similar predictions for ternary meters).

The Influence of Altered Auditory Feedback on Planning

How does altered feedback influence planning? Because models of planning have not been designed to address how planning changes in the presence of serially shifted feedback, existing models do not generate a priori predictions. One plausible prediction is that feedback triggers the production of associated events, leading to a direct relationship between distance/direction of feedback and frequency of serial-ordering errors. However, such a result may not occur, due to the fact that performers attempt to compensate for the disruptive effect of altered feedback (see Wing, 1977). Another plausible prediction that we test concerns event distinctiveness. We hypothesize that altered feedback reduces the distinctiveness of events across sequence distances because of confusions between the serial positions of planned and perceived events. We consider distinctiveness as resulting from the "processing of differences among elements that are similar on some dimension" (Hunt, 2003, p. 812; see also Brown, Preece, & Hulme, 2000; Hunt & McDaniel, 1993; Nairne, 2002).

Categorical distinctiveness arises in the range model in terms of metrical accent strength (Palmer & Pfordresher, 2003). Musical events are distinctive to the degree that they are associated with strong versus weak accents (see Equation 1). Similarity relationships are manipulated in the model in terms of metrical weights (see Equation 2). Event distinctiveness increases when higher metrical levels (Level 2 of the grid in Figure 2) receive larger weights, which leads to a reduction in weights at lower levels (Level 1 in Figure 2). The default model designates weights of .5 on each level for the metrical grid in Figure 2, generating intermediate distinctiveness. As the weight of Level 2 in Figure 2 approaches 0 (w_2 = low weight), metrical accent strengths become identical, so that no positions are dissimilar from each other on the basis of metrical accent. Alternatively, as Level 2 approaches 1 (w_2 = high weight), events become maximally dissimilar with respect to differences between strong and weak accent strengths. Intermediate weights, which we incorporate in model fits (as in Palmer & Pfordresher, 2003), produce intermediate distinctiveness. We test here whether altered auditory feedback reduces weights on Metrical Level 2 (i.e., reduced distinctiveness). We test this by fitting the range model to serial-ordering errors and comparing the model's weights on Level 2 across feedback conditions.

Three experiments, reported here, tested these issues for the first time with a novel paradigm. The first two experiments compared performance under normal auditory feedback with that under delayed feedback (Experiment 1) or prelayed feedback (Experiment 2); Experiment 3 included normal, delayed, and prelayed feedback conditions on different trials. All altered feedback presented the events as indicated in notation, termed *intended* events. We focused on AP in order to measure changes to the directional characteristics of planning and on movement error gradients (error proportions across sequence distances) in order to measure changes in the distance-related characteristics of planning due to auditory feedback.

EXPERIMENT 1
Hearing the Past

In Experiment 1, pianists performed short melodies from memory while listening to themselves over headphones. Trials incorporated a synchronization/continuation paradigm (Stevens, 1886; Wing & Kristofferson, 1973). During continuation, feedback events could originate from the current event or from previous sequence events by a lag of −1, −2, or −3 events (see Figure 2).

Method

Participants

A sample of 21 adult pianists included 16 pianists from the Ohio State University community and 5 pianists from the McGill University community (mean age = 19.9 years; range, 18–24), who participated in exchange for nominal payment or course credit in introductory psychology. The pianists had an average of 8.9 years of private piano instruction (range, 4–17) and 13.0 years of experience playing the piano (range, 6–19). Thirteen reported being right-handed; the rest were left-handed.

Materials and Apparatus

Two isochronous 12-note melodies (used also in Pfordresher, 2003) served as stimulus materials. One melody was notated in the key of G major; the other was in C major. Both melodies featured a binary (2/4) meter and were performed with the right hand. Pitches in each melody did not repeat within spans of three events, so that the serial shifts would not present the same pitch as the intended event.

The participants performed the melodies on a Roland weighted-key digital piano,[4] which simulates the feel of an acoustic piano. Presentation of auditory feedback and metronome pulses, as well as MIDI data acquisition, was implemented via the FTAP software program (Finney, 2001). The participants heard performances and metronome pulses over AKG-K270 headphones; the piano timbre originated from Program 1 (Standard Concert Piano 1), and the metronome timbre originated from Program 126 (Standard set, MIDI Key 56 = "cowbell") of the digital piano. MIDI velocity of auditory feedback was held constant; the sound intensity was approximately 80 dB SPL (measured by a General Radio Model 1982 sound-level meter centered at 1 kHz with A weighting, coupled to a TDH-39 audiometer earphone).

Design and Conditions

Each pianist performed one melody in five different auditory feedback conditions (three altered and two normal feedback). In the altered feedback conditions, each keystroke triggered a feedback pitch that matched one of the pitches in the sequence but was separated from the current (produced) position by lags of 1, 2, or 3 events (see Figure 1). In one of the normal feedback conditions, each keystroke triggered the intended (correct) event associated with the current position (lag 0); produced events differed from auditory feedback events in this condition only if a performer made an error (i.e., the keystroke still triggered the correct pitch). Auditory feedback was manipulated in this way to maintain parity between the normal and the delayed feedback conditions. In the second normal feedback condition (standard), the participants heard auditory feedback that would typically result from keypresses, including any errors that they produced.

The pianists performed eight repetitions of each feedback condition, yielding a total of 40 trials in a within-subjects design. Each participant was assigned to one of two melodies and to one of two random orders of trials. Trials were grouped into eight blocks over a single experimental session. Each block included one repetition of all five feedback conditions, presented in random order. The first trial in a session was always the standard feedback condition (in which the participants heard actual, rather than intended, auditory feedback).

Procedure

At the beginning of a session, the participants practiced the first melody with standard auditory feedback until it was memorized and performed without errors, after which the music notation was removed. Then the participants performed at least two repetitions of that melody from memory in synchrony with the metronome at the prescribed rate (500 msec between metronome onsets), without pausing between repetitions. Then the participants performed it with a lag 1 delay at a comfortable self-selected rate for another two repetitions. Following this familiarization with synchronization and delayed feedback, the participants performed at least one practice trial with the same altered feedback condition.

In the initial synchronization phase of each trial, the participants performed a melody with the metronome (500 msec between metronome onsets) and standard feedback (corresponding to keypresses). After 12 note onsets (one repetition of the melody), the metronome stopped, and the participant attempted to maintain that rate during the continuation phase while one of the five auditory feedback conditions took place. The continuation phase lasted for another 24 keypresses (two repetitions of the melody for error-free performances), at which time the cessation of feedback signaled the end of the trial.

Data Analyses

Disruption was measured in terms of pitch errors and produced timing (using interonset intervals [IOIs], the time that elapsed between two successive keypresses) in the continuation phase of each trial. Trials in which the participants made any error during synchronization were excluded if those errors resulted in an insertion of an extra event or the deletion of an event, because such errors could change the separation between produced actions and feedback events during the continuation phase. For instance, the insertion of an extra event during synchronization would effectively change a lag 1 feedback condition to a lag 2 feedback condition.

Disruption of accuracy was measured by the proportion of remaining trials in each feedback condition (out of eight repetitions) that contained any pitch error. Errors were detected with software that compared produced pitches with those that would occur in a correct performance (Large, 1993; Palmer & van de Sande, 1993, 1995). The proportion of trials with any error (*proportion of trials in error*) was used to measured accuracy, rather than error rates per trial, because errors following the first error can change the separation between produced actions and feedback events.

Timing disruption was measured by the mean and coefficient of variation (CV = *SD*/mean IOI within a trial) for IOIs; CV controls for the covariation of timing variability and production rate (Wing & Kristofferson, 1973). Events after the first produced error during continuation were discarded from timing analyses. Timing outliers (defined as ±3 *SD*s around the mean IOI) were then removed, and any possible influence of linear tempo drift was removed by adding the mean IOI for a trial to the residual IOI values from a linear regression of IOI on sequence position. On average, 20 events per trial (out of 24 possible) contributed to timing analyses across experiments.

Serial-ordering errors were analyzed as a function of distance for performances with normal versus delayed auditory feedback (averaged across feedback distance).[5] As with measures of overall accuracy, trials with deletion or addition errors during synchronization were eliminated, and only the first error produced was included in each trial. This procedure resulted in a sample of 442 errors, on average, across experiments (474, 374, and 479 for Experiments 1–3, respectively), 73% of which were serial-ordering errors (deletions, representing 16%, were the next most common). Only errors that matched a pitch within distances of three events—the range over which auditory feedback manipulations occurred—were examined (72% of the serial-ordering errors). Missing data were replaced with means across the other participants for that feedback condition.

Results

Preliminary analyses indicated that the lag 0 and standard feedback conditions did not differ from each other in measures of disruption; therefore, all the analyses used the lag 0 condition (in which performers heard intended events) as representative of *normal* feedback, because it best matched the altered feedback conditions.

Overall Disruption

Two one-way ANOVAs were conducted on each measure of disruption. The first ANOVA tested overall disruption by comparing average disruption across altered feedback conditions (lags of 1–3) with normal feedback. The second ANOVA compared disruption across altered feedback conditions.

The mean proportion of trials in error, shown in Figure 3, revealed significant disruption from delayed feedback that did not vary with feedback distance. There was a significant main effect of feedback type [normal or altered; $F(1,20) = 24.32$, $MS_e = 0.026$, $p < .01$]. There were no differences among different delay lags in the altered feedback conditions ($p < .10$). The results for measures of timing (mean IOI, CV) mirrored those found for errors. Produced IOIs were longer [$F(1,17) = 5.59$, $MS_e = 55.13$, $p < .05$] and were more variable [higher CVs; $F(1,17) = 8.65$, $MS_e < 0.001$, $p < .01$] for performances with delayed than for those with normal feedback. The amount of timing disruption also did not differ across feedback distances. The effect size was larger for errors than for the timing measures (partial ω^2 for errors = .36; for CV, IOI = .14; for MN, IOI = .10; Keppel, 1991, p. 354).

Distance Effects in Serial-Ordering Errors

Obtained data. Serial-ordering errors were analyzed as a function of distance for performances with normal versus delayed auditory feedback (averaged across feedback distance). Because serial-ordering errors in the altered feedback conditions did not differ reliably across delay lags, movement gradients were averaged across all the altered feedback conditions. Figure 4 shows movement gradients (mean proportions of serial-ordering errors originating from sequence distances 1–3) for normal (left) and altered (right) feedback conditions. Chance estimates for errors, indicated by the dashed line, display the proportions of errors that could originate from distances 1–3 on the basis of the number of produced events in a trial.[6]

As is shown in Figure 4, most errors originated from distances of 2 (metrically similar events), and the fewest errors originated from distances of 3 (metrically dissimilar and less proximal events), for both the normal and the altered feedback conditions, consistent with the predictions of the range model (Palmer & Pfordresher, 2003). A two-way ANOVA on error proportions by sequence distance (1–3) and feedback condition (normal or altered) revealed a significant main effect of distance [$F(2,40) = 4.48$, $MS_e = 0.115$, $p < .05$] but no interaction with feedback condition ($F < 1$).

Model fits. Fitted weights on the second level of the metrical grid were computed for the normal and the altered feedback conditions. First, the range model was fit to the data of individual participants, with equal weights on each metrical level (w_1, $w_2 = .50$; see Equation 2), to generate estimates of the initial activation parameter a for each participant. The a parameter was allowed to vary between .8 and .99 in increments of .001 (as in Palmer & Pfordresher, 2003); the mean optimal a was .913. The model was then fit again with a fixed while weights on Level 2 (w_2) were allowed to vary between .01 and .99, in steps of .01. Fits to the normal feedback conditions (median variance accounted for [VAF] = .91) were generally better than fits to the altered feedback conditions (median VAF = .67), although this difference did not reach significance. Of most interest, fitted weights at Level 2 of the metrical hierarchy were significantly larger for the normal feedback conditions (median $w_2 = .64$) than for the delayed conditions [$w_2 = .47$; Wilcoxon $T(21) = 19.0$, $p < .01$].

Direction Effects in Serial-Ordering Errors

The AP was computed for the same subset of errors as that used to measure the distance effects. The AP was calculated for each participant and condition by dividing the number of anticipatory errors by the sum of anticipatory and perseveratory errors. Table 1 shows the mean AP for each feedback condition in all the experiments. AP increased when the participants heard delayed feedback, relative to normal feedback. A one-way ANOVA that compared AP for normal versus altered feedback yielded a main effect of feedback condition [$F(1,20) = 6.62$, $MS_e = 0.052$, $p < .05$]. A second ANOVA that analyzed AP across the altered feedback conditions yielded a main effect of feedback distance [$F(2,40) = 8.14$, $MS_e = 0.030$, $p < .01$]. Post hoc tests verified that AP was lower for feedback distances of 2 than for other feedback distances. Finally, a planned comparison assessed the difference between normal and lag 1 feedback, given the assumption that postoutput suppression (which presumably influences the directional characteristics of planning) operates primarily on the most recently produced event. AP was significantly higher for lag 1 than for normal feedback [$t(20) = 3.83$, $p < .01$]. Mean frequencies of anticipations versus perseverations suggested that change in AP occurred primarily due to decreases in perseveratory errors: Perseverations decreased 59% in the altered feedback conditions (relative to the normal ones), whereas anticipations increased by 7%.

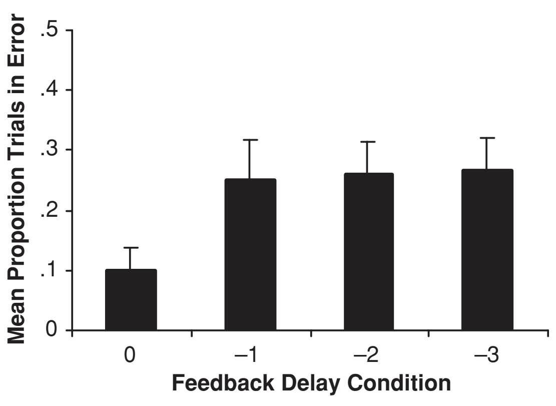

Figure 3. Mean proportions of trials in Experiment 1 that contained error(s) by feedback condition (0 = normal feedback), with standard error bars (+1 *SE*).

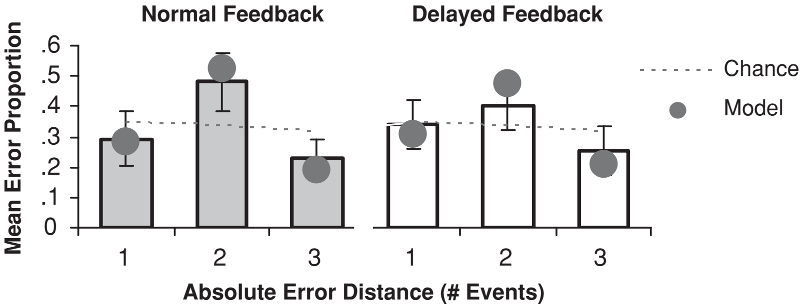

Figure 4. Mean proportions of errors by sequence distance and feedback condition for Experiment 1 (±1 *SE*), with range model predictions and chance estimates.

Discussion

Experiment 1 documented three effects of hearing the past on music performance. Delayed feedback significantly disrupted accuracy but caused less disruption to timing, consistent with the results in Pfordresher (2003). Second, fits of the range model to serial-ordering errors suggested that event distinctiveness decreased in the presence of delayed feedback, relative to normal feedback. This was seen in lowered metrical weights at higher hierarchical levels, which decreases the salience of strong beats relative to weak beats, under altered feedback. Third, hearing the past increased the proportion of anticipatory errors, relative to normal feedback, even when overall error rates increased, contrary to the general anticipatory effect (Dell et al., 1997). Thus, Experiment 1 supports the hypothesis that performers alter planning to compensate for delayed feedback. We further pursued this compensatory explanation under prelayed feedback conditions in the next experiment.

EXPERIMENT 2
Hearing the Future

Experiment 2 explored a new manipulation of auditory feedback: Pianists heard events intended for the future in altered feedback conditions. On different trials, feedback events originated from the current event or from future sequence events, separated by leads of +1, +2, or +3 events (see Figure 1). On the basis of the findings in Experiment 1 (delayed feedback increased AP), we predicted that prelays would cause a compensatory shift in planning in the direction opposite to the feedback, leading to a decrease in AP, relative to normal feedback conditions.

Method

Participants

A sample of 21 adult pianists included 16 pianists from the Ohio State University community and 5 pianists from the McGill University community (mean age = 24.2 years; range, 18–46), who participated in exchange for nominal payment or course credit in introductory psychology. The pianists had an average of 10.1 years of private piano instruction (range, 2–20) and 15.8 years of experience playing the piano (range, 5–44). Sixteen reported being right-handed; the rest were left-handed. None had participated in the previous experiment.

Materials, Conditions, and Procedure

The materials were identical to those used in Experiment 1. The design and procedure were also identical, except that the altered feedback conditions presented prelays of lead +1, +2, or +3 during the continuation phase, which resulted in the participants' hearing pitches associated with events from the future. The two normal feedback conditions were identical to those used in Experiment 1.

Results

Overall Disruption

As in Experiment 1, no difference was found between the two normal feedback conditions, and *normal* feedback henceforth will refer to the condition in which feedback events matched the current intended event. Mean proportions of trials in error are shown in Figure 5 across the feedback conditions. Prelayed feedback significantly increased errors, relative to normal feedback, irrespective of the amount of lead. This was confirmed by a main ef-

Table 1
Mean Anticipatory Proportions of Errors Across Feedback Conditions and Experiments

	Delays					Prelays			
	M	−3	−2	−1	Norm	±1	±2	±3	*M*
Experiment 1	**.88**	.96	.76	.94	.70				
Experiment 2					.85	.64	.82	.80	**.75**
Experiment 3	**.72**	.71	.72	.73	.89	.74	.82	.81	**.79**
Reversals	.67	.69	.61	.70		.77	.86	.88	.81
Nonreversals	**.84**	.86	.77	.88		.73	.59	.78	**.73**

Note—Boldface highlights means from comparable altered feedback conditions. See the text for details.

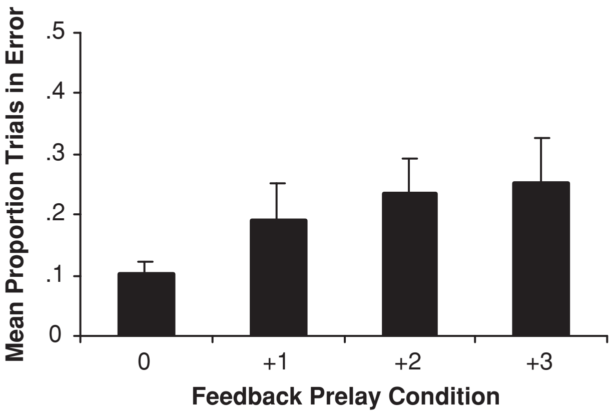

Figure 5. Mean proportions of trials in Experiment 2 that contained error(s), by feedback condition (0 = normal feedback), with standard error bars (+1 *SE*).

fect of feedback condition in the first ANOVA, in which normal feedback was compared with altered feedback [$F(1,21) = 7.78$, $MS_e = 0.035$, $p < .05$]. The second ANOVA revealed no differences across amounts of prelay ($F < 1$). Analyses of timing did not reveal disruption from altered feedback, relative to normal feedback. Follow-up analyses compared proportions of trials in error across Experiments 1 and 2; these analyses revealed no differences as a function of experiment.

Distance Effects in Serial-Ordering Errors

Obtained data. Figure 6 shows movement gradients for serial-ordering errors in normal feedback conditions (left) and altered feedback conditions (right). One participant, who generated no serial-ordering errors, was excluded from these analyses. As is predicted by the range model (Palmer & Pfordresher, 2003), errors were more likely to originate from metrically similar events at a distance of 2 (to an even larger degree than was found in Experiment 1), and errors from the farthest distance of 3 were least prevalent. More important, the peak at a distance of 2, evidenced again in normal feedback trials, vanished during trials with prelayed feedback. These two findings were confirmed by the significant error distance × feedback condition interaction for error proportions [$F(2,40) = 4.26$, $MS_e = 0.077$, $p < .05$] and a main effect of error distance [$F(2,40) = 20.76$, $MS_e = 0.077$, $p < .01$]. A follow-up ANOVA included the error data from Experiment 1 and added the factor of feedback direction (delayed and prelayed), which did not interact with any of the factors.

Model fits. Model fits for metrical weights on Level 2 were conducted as before. The mean optimal *a* across individual fits was .91. Fits were better for normal (median VAF = 1.00) than for prelayed (median VAF = .78) feedback conditions, although this difference did not reach significance. As in Experiment 1, the Level 2 weights were significantly larger for normal feedback conditions (median $w_2 = .85$) than for prelayed feedback (median $w_2 = .58$) [Wilcoxon $T(19) = 36$, $p < .01$]. The reduced weights for prelayed feedback approached the baseline value of .5 (an absence of weighting). Thus, strong and weak beats were less distinctive in prelayed feedback trials.

Obtained metrical weight parameters (w_2) were also compared across experiments. Values of w_2 did not differ for delays (Experiment 1) versus prelays (Experiment 2; $p > .10$). However, w_2 did differ across experiments for the normal feedback conditions (Mann–Whitney $U = 121.5$, $p < .01$), with higher weights for the normal feedback conditions in Experiment 2 than for those in Experiment 1. Thus, delays and prelays appeared to affect event activations as a function of feedback direction, but only in terms of a carryover effect to normal feedback trials.

Direction Effects in Serial-Ordering Errors

Mean AP values for Experiment 2 are shown in Table 1. The influence of feedback direction on AP was opposite to the result from Experiment 1: AP was lower in the prelayed feedback ($M = .75$) than in the normal feedback conditions ($M = .85$). The ANOVA comparing normal with altered feedback approached significance ($p = .07$). The ANOVA comparing AP across the altered feedback conditions yielded a significant effect of feedback distance [$F(2,40) = 3.53$, $MS_e = 0.057$, $p < .05$]. Post hoc tests (Tukey's HSD = .05) verified that prelays with a distance

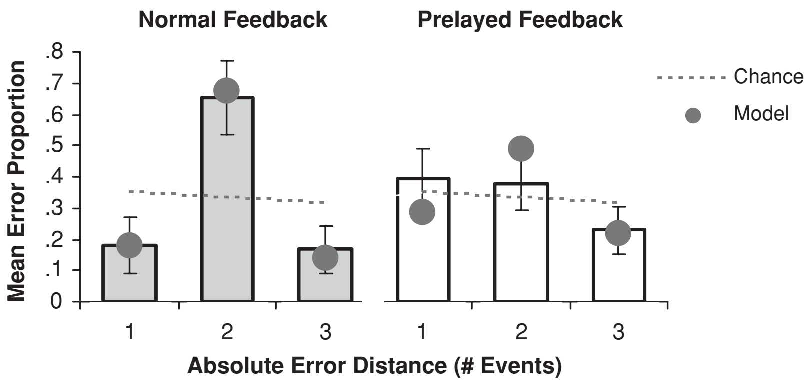

Figure 6. Mean proportions of errors by sequence distance and feedback condition for Experiment 2 (±1 *SE*), with range model predictions and chance estimates.

of 1 elicited lower AP than did prelays from farther distances. Finally, the planned comparison between the normal and the lead 1 prelay condition yielded a significant effect [$t(20) = -3.25, p < .01$]. Change in AP occurred primarily due to increases in perseveratory errors: Perseverations increased 81% in the prelay conditions, relative to normal feedback, whereas anticipations increased only by 66%.

The influence of feedback direction on AP was assessed by comparing AP from Experiments 1 and 2. The effects of delayed and prelayed feedback from a distance of 1 (of most interest in explanations of postoutput suppression) were compared with a 2 (feedback type: normal or altered) × 2 (feedback direction: delayed or prelayed) ANOVA that yielded a significant feedback type × direction interaction [$F(1,40) = 25.16$, $MS_e = 0.045$, $p < .01$] and no main effects. Post hoc tests (Tukey's HSD, $\alpha = .05$) verified that Experiment 1 produced lower AP for normal than for delayed feedback, whereas Experiment 2 produced higher AP for normal than for prelayed feedback.

Discussion

Experiment 2 explored the effects of a new kind of feedback alteration, prelays, in which performers heard events intended for the future during performance. As with delays, hearing the future increased errors and flattened the error proportions arising from different sequence distances, which suggests reduced distinctiveness. Prelays influenced the directional characteristics of planning in a way opposite to that for delays by decreasing the anticipatory tendency, relative to normal feedback. This pattern is consistent with a compensatory planning explanation.

Thus, Experiments 1 and 2 show that hearing the past influences direction of planning (AP) differently than does hearing the future, whereas both feedback directions yield similar influences on distance relationships in planning. Some of these findings may have been influenced by a session effect. Feedback direction was always consistent across the altered feedback trials in Experiments 1 and 2. In Experiment 3, we examined the effect of directional variability by presenting delays or prelays on different trials in an unpredictable order.

EXPERIMENT 3
Hearing the Past or Future

The participants in Experiment 3 experienced delayed or prelayed feedback on different trials, which were randomly ordered with normal feedback trials. The session was structured so that changes in feedback type across successive trials were likely (86% of the trials). The six altered feedback conditions included delays (lags -1, -2, and -3) and prelays (leads $+1$, $+2$, and $+3$) and a single normal feedback condition that presented the intended current event (lag/lead 0).

Method

Participants

A sample of 28 adult pianists included 12 pianists from the Ohio State University community and 16 pianists from the University of Texas at San Antonio community (mean age = 23.0 years, range, 16–57), who participated in exchange for course credit in an introductory psychology or nominal payment. The pianists had an average of 9.9 years of private piano instruction (range, 4–30) and 17.1 years of experience playing the piano (range, 6–53). Twenty-one reported being right-handed, 1 reported being ambidextrous, and the rest reported being left-handed. None had participated in the previous experiments.

Materials

The same materials were used as those in Experiments 1 and 2. The participants performed both stimulus melodies in all the experimental conditions.

Design

Seven auditory feedback conditions were included; three conditions included different feedback delays (lag of -1, -2, or -3), and three included different feedback prelays (lead of $+1$, $+2$, or $+3$). The seventh feedback condition was a normal feedback condition in which the performers heard the correct (intended) sequence events. Within-subjects variables for Experiment 3 included feedback condition (seven levels), repetition (five), and stimulus melody (two), yielding 70 trials in a session. The participants performed a different melody in each half of the session. Half-sessions were further divided into five blocks, and the participants experienced each of the seven feedback conditions, randomly ordered, within each block. The first trial in each half session was always from the normal feedback condition. The orders of the two stimulus melodies and two random orders of trials were counterbalanced across participants. Of the trials, 86% featured either a reversal in feedback direction from the previous trial or a change from normal feedback to a delay or prelay.

Procedure

The procedure was identical to that used in Experiments 1 and 2, except that there was a break halfway through the session, after which the participants learned and memorized a new melody.

Results

Overall Disruption

The mean proportions of trials in error are shown in Figure 7. The first ANOVA on overall disruption yielded a significant main effect of feedback type [normal, prelay, and delay; $F(2,54) = 7.88$, $MS_e = 0.019$, $p < .01$]. Post hoc comparisons (Tukey's HSD, $\alpha = .05$) confirmed that both delays ($M = .355$) and prelays ($M = .348$) elicited higher disruption than did normal feedback ($M = .225$) but did not differ from each other. The second ANOVA, which examined the interaction of feedback direction (delay or prelay) with feedback distance (± 1, 2, or 3), revealed no significant effects. Timing measurements (mean IOIs, CVs) in Experiment 3 revealed no disruptions from delays or prelays, relative to normal feedback, or effects related to the distance of altered feedback. Follow-up analyses that compared proportions of trials in error in Experiment 3 with those in comparable conditions in the earlier experiments revealed no effects as a function of experiment.

Distance Effects in Serial-Ordering Errors

Obtained data. Mean movement gradients are shown in Figure 8 for normal (middle), delayed (left), and prelayed (right) conditions in Experiment 3. Performances with normal and delayed feedback fit the qualitative pre-

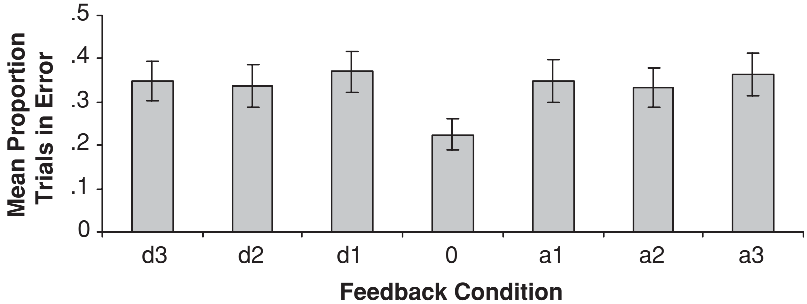

Figure 7. Mean proportions of trials in Experiment 3 that contained error(s) by feedback condition (0 = normal feedback), with standard error bars (+1 *SE*).

dictions of the range model, with higher error proportions originating from metrically similar events (at a distance of 2). Performances with prelayed feedback do not show a peak at an error distance of 2 (as in Experiment 2), which suggests that event distinctiveness from metrical similarity was reduced in these performances. Two ANOVAs were run that matched the analyses used in the earlier experiments. Each used a 2 (feedback condition) × 3 (error distance) design. The first ANOVA compared normal and delayed feedback and yielded a main effect of error distance [$F(2,54) = 10.62$, $MS_e = 0.140$, $p < .01$] but no interaction. The second ANOVA compared normal and prelayed feedback and likewise yielded a main effect of error distance [$F(2,54) = 10.19$, $MS_e = 0.126$, $p < .01$] but no interaction, despite the fact that prelays elicited a qualitatively different movement gradient than did normal feedback. A third ANOVA compared error profiles across the two altered feedback conditions within Experiment 3. This ANOVA led to both a main effect of error distance [$F(2,54) = 15.78$, $MS_e = 0.298$, $p < .01$] and a significant error distance × feedback direction interaction [$F(2,54) = 3.80$, $MS_e = 0.123$, $p < .03$; familywise $\alpha = .05$, modified Bonferroni correction; Keppel, 1991, p. 169]. Follow-up analyses that compared movement gradients from Experiment 3 with those in comparable conditions in the earlier experiments revealed no effects as a function of experiment.

Model fits. Model fits for metrical weights on Level 2 were conducted as in Experiment 1. The average optimal *a* was .88, similar to those in the other two experiments. Fits were better in the normal feedback conditions (median VAF = .87) than in the delayed (median = .80) or prelayed (median = .67) feedback conditions, although these differences were not significant. Of most interest, the weight for normal feedback (median $w_2 = .63$) was similar to that for delayed feedback (median $w_2 = .67$) but was higher than that for prelayed feedback (median $w_2 = .45$) [Wilcoxon $T(28) = 112.0$, $p < .05$]. Likewise, the weight was higher for delayed than for prelayed feedback [Wilcoxon $T(28) = 91.0$, $p < .01$]. Thus, strong and weak beats were less distinctive in the prelayed feedback trials than in the other conditions.

Direction Effects in Serial-Ordering Errors

All altered feedback conditions reduced AP ($M = .75$), relative to normal feedback ($M = .89$). An ANOVA comparing AP across the normal, delayed, and prelayed conditions (averaged across feedback distance) yielded a main

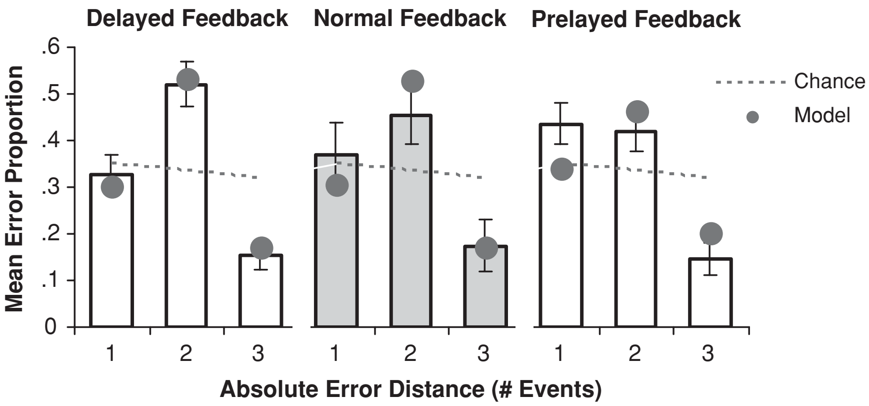

Figure 8. Mean proportions of errors by sequence distance and feedback condition in Experiment 3 (±1 *SE*), with range model predictions and chance estimates.

effect of feedback condition [$F(2,54) = 6.59$, $MS_e = 0.030$, $p < .05$]. Table 1 shows the mean AP values for each feedback condition. A second ANOVA that compared the influences of feedback distance and direction yielded no main effects or interaction. Planned comparisons between the normal feedback conditions and lags or leads from a distance of 1 revealed significant effects for prelays [$t(27) = -2.37$, $p < .05$], as well as for delays [$t(27) = 3.06$, $p < .01$]. As in Experiments 1 and 2, change in AP was related mostly to perseveratory errors; perseverations increased by 73% in the delayed feedback trials (anticipations increased by 21%) and by 44% in the prelayed feedback trials (anticipations increased by 44%).

The effect of trial-to-trial variability in feedback conditions was further addressed by categorizing altered feedback trials into two groups on the basis of whether the previous trial presented the opposite feedback direction (e.g., from delay to prelay) or not (e.g., two successive delayed trials, or from normal to delayed). As is shown in Table 1, trials that did not follow a reversal showed a pattern that was qualitatively the same as that seen in Experiments 1 and 2: AP was higher for delays ($M = .84$) than for prelays ($M = .70$). In contrast, trials that followed a reversal in feedback direction showed a pattern opposite to that found in Experiments 1 and 2, with AP higher for prelays ($M = .83$) than for delays ($M = .67$). This observation was verified in an ANOVA that included the factors of trial type (reversal or no reversal) and feedback direction (delays or prelays), which yielded a significant interaction [$F(1,27) = 33.13$, $MS_e = 0.036$, $p < .01$]. Performers compensated for the direction of feedback in their direction of planning as in the earlier experiments, when that feedback direction was consistent across trials. The same pattern of results emerged when the analysis excluded all trials that followed normal feedback trials.

Discussion

Experiment 3 replicated the main findings of overall disruption and serial-ordering error distributions in Experiments 1 and 2 when delayed and prelayed conditions were included within the same experiment. One difference, related to the change in feedback direction across trials, was noted: AP was reduced, relative to normal feedback, for all the delay conditions. This difference resulted from changes in feedback direction across trials. The results for AP resembled those in Experiments 1 and 2 when feedback direction was consistent across successive trials, whereas the results for AP were opposite to the results in Experiments 1 and 2 when feedback direction reversed. Thus, AP is sensitive to the consistency of altered feedback across successive trials.

GENERAL DISCUSSION

In three experiments, we examined the production of musical sequences when pitches of auditory feedback events were altered to match events planned for other sequence positions. Relationships between produced events and auditory feedback were varied with respect to direction (past or future) and distance (serial separation in past or future). Three main conclusions emerged. First, delayed and prelayed auditory feedback disrupted the accuracy of produced events but spared their timing (cf. Pfordresher, 2003). Overall accuracy was influenced to a similar degree, regardless of feedback direction or distance. Second, both delays and prelays disrupted event distinctiveness in the planning of serial order by equalizing the metrical accent strength of events, relative to performance during normal feedback (cf. Palmer & Pfordresher, 2003). Third, AP varied reliably as a function of feedback distance and direction. In particular, performers shifted the direction of planning in a compensatory fashion in response to the direction of auditory feedback, although this compensation was apparently contingent on the regularity with which altered feedback conformed to one direction or another.

We have proposed an account of feedback disruption that is based on matches between planned and perceived events. Performers may respond to perception–action similarity on a global level that relates two sequences—those from auditory feedback and planned serial order—rather than respond to local matches between individual planned and perceived events. This view follows from the range model of planning (Palmer & Pfordresher, 2003), in which each event activation is stored as a contextually defined representation of an entire planning increment. The idea that perception–action similarity relationships rely on the global organization of events explains why performers are disrupted by serial shifts, but not by feedback sequences that contain random permutations of pitches from produced melodies (Finney, 1997; Pfordresher, 2005).

Auditory Feedback and Planning

Disruption occurred when feedback contents were altered to resemble planned events. We suggest that disruption occurs because feedback adds activation to events within the planning increment, resulting in interference when feedback matches events planned for noncurrent events. Changes to planning do not simply reflect the sum of the original planned activations and the added activation from auditory feedback. Instead, changes to planning caused by altered feedback occur because performers try to shift activation away from events associated with feedback alterations. Performers shift direction of planning in a compensatory manner, yielding changes in AP and reductions in categorical event distinctiveness. The fact that error rates were equivalent across the altered feedback conditions suggests that the performers experienced disruption in all the conditions.

Compensatory responses were apparent in directional characteristics of planning (AP). Delays increased AP, relative to normal feedback, in Experiment 1, whereas prelays reduced AP in Experiment 2. Experiment 3, with both forms of altered auditory feedback, was particularly diagnostic. The performers appeared to maintain the directional orientation of the previous trial. When successive trials did not reverse in feedback direction, the results resembled the findings in Experiments 1 and 2. In contrast, when feedback direction reversed across trials, AP varied in the opposite direction and was higher for prelays than for delays. Thus, changes to direction in planning may take time and may accumulate over multiple altered feedback trials that present a consistent feedback direction.

Altered feedback conditions generally reduced the tendency in the normal feedback conditions for errors to originate from the most metrically similar events, consistent with previous findings (Drake & Palmer, 2000) and with the range model of planning (Palmer & Pfordresher, 2003). This pattern of results suggests that altered auditory feedback reduces metrical distinctiveness among planned events, consistent with the characterization of metrical accent as a categorical dimension of musical events (e.g., Palmer & Krumhansl, 1990). In addition, the range model's fit of metrical weights to serial-ordering errors that reflected metrically similar events at different hierarchical levels consistently indicated higher weights in normal feedback conditions than in altered feedback conditions. Almost all movement error gradients revealed a reduction in error confusions among metrically similar events under altered feedback.

Comparisons With Alternative Approaches

Two theories concerning the role of auditory feedback in sequence production contrast with this perspective. One view (the EXPLAN model; Howell, 2004a, 2004b; Howell & Au-Yeung, 2002) states that altered auditory feedback disrupts execution (related to timing) but does not disrupt planning (related to accuracy). The three experiments described here demonstrated disruption of production accuracy, but not of timing, when only the pitch characteristics of auditory feedback were altered and feedback was synchronized with actions. The present results also differ from the predictions of MacKay's (1987) node structure theory. Node structure theory was designed primarily as a model of speech production, with the assumption that other behaviors, such as music performance, would follow similar principles. MacKay's (1987) theory suggests that disruption from altered feedback occurs when a feedback event sounds during a content node's hypersensitive phase (~200 msec after an event has been produced). Our data present an exception to this prediction, in that we found disruption for delays of feedback that occurred over a longer time scale (500–1,500 msec) than node structure theory predicts and we found disruption for feedback that presented future as well as past events. According to node structure theory, planning of production should be immune to feedback alterations of this sort, due to the dynamics of content nodes.

Our conclusions about the role of anticipatory planning (AP) also differ from aspects of Dell et al.'s (1997) model of serial order in sentence production. In that model, AP represents a characteristic of planning that enhances fluency; the model predicts an increase in AP with practice, with slowing of production rate and with an increased anticipatory activation parameter. In contrast, the present experiments showed increased error rates that were accompanied by increased AP in the presence of altered feedback. Thus, increased AP may not always signal increased fluency but may reflect strategic changes to planning brought about by disruptive situations, such as altered feedback.

Although Dell et al.'s (1997) model was not designed to address the role of auditory feedback, some modifications of their model might account for these findings. One possibility is that feedback alters the function of a turn-off mechanism; in Dell et al., a parameter controls residual activation from the past that results from its turn-off. Increased residual activation, which might be expected from reactivation of past events by delayed feedback, leads to more perseverations in this model, relative to anticipations. This prediction is not consistent with the findings reported here; AP actually increased under delayed feedback. Other possibilities are that feedback alters anticipatory activation or a decision variable, both of which are fixed parameters in Dell et al. In the present experiments, the participants were able to reverse the bias toward future or past events within an experimental session, suggesting a variable parameter. Further evidence is needed to distinguish between process variables that can account for compensatory weighting of the future or the past in response to altered feedback.

A possible limitation of the present experiments is generalizability from memorized musical materials. Although the use of memorized sequences allows us to eliminate the possibly confounding influence of eye movements during reading that may be influenced by other notated events that are heard in altered feedback, performance from memory necessarily raises the memory load and, perhaps, the degree of disruption. Nevertheless, disruption from altered feedback is not unique to music; the influence of altered feedback disrupts many verbal and manual production tasks (Chase, Harvey, Standfast, Rapin, & Sutton, 1959; Chase, Sutton, & Rapin, 1961; Howell et al., 1983; Roberts & Gregory, 1973), as well as the production of both spontaneous and scripted sequences (Collins & Worthington, 1978; MacKay, 1968). Thus, it seems likely that altered feedback of the kind used here will yield similar effects across other tasks and stimulus materials.

Consideration of Alternative Similarity Metrics

The range model measures similarity relationships among events on the basis of metrical accent strengths. Metrical accent is only one basis on which to gauge the similarity/distinctiveness of planned musical events; other characteristics include tonality, rhythm, and timbre (Gabrielsson, 1973; Krumhansl, 1990; N. A. Smith & Schmuckler, 2004). Meter is advantageous because it offers explanatory power for the majority of serial-ordering errors (Drake & Palmer, 2000; Palmer & Pfordresher, 2003), and it offers a computationally simple framework similar to slot-filler mechanisms that can explain planning of serial order (e.g., Shattuck-Hufnagel, 1979).

Nevertheless, we will briefly consider here two possible alternative similarity metrics. One possible metric compares pitches of events separated by different distances as a function of their tonal status (related to chroma), as defined by past work on the role of tonality in music perception (e.g., Krumhansl, 1990; Krumhansl & Kessler, 1982; Krumhansl & Shepard, 1979). To test this possibility, tonal similarity relationships among pitches were tested by inserting values, obtained by Krumhansl and Kessler (1982), into Equation 1 for the stimulus melodies. The resulting similarity predictions, based on normal feedback conditions, differed from the predictions of metrical similarity. Whereas metrical similarity relationships generated a peak at a distance of 2 and equal values for distances of 1 and 3 (similarity values = .286, .429, and .286 for distances of 1–3, respectively), tonal similarity relationships predicted a decrease in similarity across distances of 1–3 (similarity values = .836, .811, and .643 for distances of 1–3).

We also considered a possible role of the fingering used to produce tones in accounting for similarity relationships, on the basis of previous work that has linked accuracy in manual sequence production to ergonomic constraints (e.g., for piano, Parncutt, Sloboda, Clarke, Raekallio, & Desain, 1997, and Sloboda, Clarke, Parncutt, & Raekallio, 1998; for typing, see Rumelhart & Norman, 1982). Pitches were coded by fingers used to produce them (1–5) on the basis of the prescribed fingering in the music notation given to participants (see Pfordresher, 2003, for stimulus fingering). Fingering codes were then entered into Equation 1 to generate similarity relationships for the melodies. As with tonal similarity, finger similarity values differed from metrical similarity predictions based on normal feedback conditions and generated a linearly decreasing series (similarity values = .630, .605, and .542). Thus, metrical similarity is unique among these indices in predicting peak similarity at a distance of 2. This result is significant because other similarity metrics could not predict the patterns of serial-ordering errors found for normal feedback conditions in the present study, without substantial changes to the underlying model.

CONCLUSION

Music performance is disrupted when performers hear events associated with future or past events. These alterations of auditory feedback disrupt the planning of serial order, rather than timing. Hearing events from the past or the future during production caused an overall reduction of distinctiveness among planned events, as observed in serial-ordering errors. Reductions in distinctiveness were accompanied by compensatory changes in the direction of planning; delays and prelays increased the tendency to produce events in the direction opposite to the feedback (evidenced in the anticipatory proportion of errors). These findings support an integrated approach to perception and action that guides retrieval in production: Perceiving one's own correct productions can reinforce planned events, and perceiving related (altered) productions influences event planning.

REFERENCES

Black, J. W. (1951). The effect of delayed side-tone upon vocal rate and intensity. *Journal of Speech & Hearing Disorders*, **16**, 56-60.

Borden, G. J. (1979). An interpretation of research on feedback interruption in speech. *Brain & Language*, **7**, 307-319.

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, **107**, 127-181.

Chase, R. A. (1965). An information-flow model of the organization of motor activity. I: Transduction, transmission and central control of sensory information. *Journal of Nervous & Mental Disease*, **140**, 239-251.

Chase, R. A., Harvey, S., Standfast, S., Rapin, I., & Sutton, S. (1959). Comparison of the effects of delayed auditory feedback on speech and key tapping. *Science*, **129**, 903-904.

Chase, R. A., Sutton, S., & Rapin, I. (1961). Sensory feedback influences on motor performance. *Journal of Auditory Research*, **3**, 212-223.

Collins, J. K., & Worthington, A. G. (1978). Delayed auditory feedback and cognitive complexity of tasks. *American Journal of Psychology*, **91**, 93-99.

Conrad, R. (1965). Order error in immediate recall of sequences. *Journal of Verbal Learning & Verbal Behavior*, **4**, 161-169.

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

Cooper, R., & Shallice, T. (2000). Contention scheduling and the control of routine activities. *Cognitive Neuropsychology*, **17**, 297-338.

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

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

Drake, C., & Palmer, C. (2000). Skill acquisition in music performance: Relations between planning and temporal control. *Cognition*, **74**, 1-32.

Duke, R. A. (1989). Musicians' perception of beat in monotonic stimuli. *Journal of Research in Music Education*, **37**, 61-71.

Fairbanks, G. (1954). Systematic research in experimental phonetics: 1. A theory of the speech mechanism as a servosystem. *Journal of Speech & Hearing Disorders*, **19**, 133-139.

Fairbanks, G., & Guttman, N. (1958). Effects of delayed auditory feedback upon articulation. *Journal of Speech & Hearing Research*, **1**, 12-22.

Finney, S. A. (1997). Auditory feedback and musical keyboard performance. *Music Perception*, **15**, 153-174.

Finney, S. A. (1999). *Disruptive effects of delayed auditory feedback on motor sequencing*. Unpublished doctoral dissertation, Brown University.

Finney, S. A. (2001). FTAP: A Linux-based program for tapping and music experiments. *Behavior Research Methods, Instruments, & Computers*, **33**, 65-72.

Finney, S. A., & Palmer, C. (2003). Auditory feedback and memory for music performance: Sound evidence for an encoding effect. *Memory & Cognition*, **31**, 51-64.

Fromkin, V. A. (1971). The nonanomalous nature of anomalous utterances. *Language*, **47**, 27-52.

Gabrielsson, A. (1973). Similarity ratings and dimension analyses of auditory rhythm patterns: I. *Scandinavian Journal of Psychology*, **14**, 138-160.

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. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Speech and talk* (pp. 177-220). London: Academic Press.

Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideo-motor mechanism. *Psychological Review*, **77**, 73-99.

Healy, A. F. (1974). Separating item from order information in short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **13**, 644-655.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral & Brain Sciences*, **24**, 849-937.

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science*, **279**, 1213-1216.

Howell, P. (2004a). Assessment of some contemporary theories of stuttering that apply to spontaneous speech. *Contemporary Issues in Communicative Sciences & Disorders*, **39**, 122-139.

Howell, P. (2004b). Effects of delayed auditory feedback and frequency-shifted feedback on speech control and some potentials for future development of prosthetic aids for stammering. *Stammering Research*, **1**, 31-77.

Howell, P., & Archer, A. (1984). Susceptibility to the effects of delayed auditory feedback. *Perception & Psychophysics*, **36**, 296-302.

Howell, P., & Au-Yeung, J. (2002). The EXPLAN theory of fluency control applied to the diagnosis of stuttering. In E. Fava (Ed.), *Clinical linguistics: Theory and applications in speech pathology and therapy* (pp. 75-94). Amsterdam: Benjamins.

Howell, P., Powell, D. J., & Khan, I. (1983). Amplitude contour of the delayed signal and interference in delayed auditory feedback tasks. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 772-784.

Hunt, R. R. (2003). Two contributions of distinctive processing to accurate memory. *Journal of Memory & Language*, **48**, 811-825.

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory & Language*, **32**, 421-445.

James, W. (1890). *Principles of psychology* (Vol. 2). New York: Holt.

Kalinowski, J., & Saltuklaroglu, T. (2003). Speaking with a mirror: Engagement of mirror neurons via choral speech and its derivatives induces stuttering inhibition. *Medical Hypotheses*, **60**, 538-543.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, **11**, 201-258.

Keppel, G. (1991). *Design and analysis: A researcher's handbook*. Upper Saddle River, NJ: Prentice-Hall.

Krumhansl, C. L. (1990). *The cognitive foundations of musical pitch*. New York: Oxford.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, **89**, 334-368.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 579-594.

Large, E. W. (1993). Dynamic programming for the analysis of serial behaviors. *Behavior Research Methods, Instruments, & Computers*, **25**, 238-241.

Lashley, K. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112-146). New York: Wiley.

Lee, B. S. (1950). Some effects of side-tone delay. *Journal of the Acoustical Society of America*, **22**, 639-640.

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

Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, **8**, 249-336.

MacKay, D. G. (1968). Metamorphosis of a critical interval: Age-linked changes in the delay in auditory feedback that produces maximal disruption of speech. *Journal of the Acoustical Society of America*, **43**, 811-820.

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

Mates, J., & Aschersleben, G. (2000). Sensorimotor synchronization: The impact of temporally displaced auditory feedback. *Acta Psychologia*, **104**, 29-44.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart & Winston.

Müller, K., Aschersleben, G., Esser, K. H., & Müsseler, J. (2000). Effects of delayed auditory feedback on speech: Just a problem of displaced rhythm? In P. Desain & L. Windsor (Eds.), *Rhythm perception and production* (pp. 249-261). Lisse: Swets & Zeitlinger.

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology*, **53**, 53-81.

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

Palmer, C., & Pfordresher, P. Q. (2003). Incremental planning in sequence production. *Psychological Review*, **110**, 683-712.

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

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

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, **11**, 409-464.

Parncutt, R., Sloboda, J. A., Clarke, E. F., Raekallio, M., & Desain, P. (1997). An ergonomic model of keyboard fingering for melodic fragments. *Music Perception*, **14**, 341-382.

Pfordresher, P. Q. (2003). Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing. *Journal of Experimental Psychology: Human Perception & Performance*, **29**, 949-964.

Pfordresher, P. Q. (2005). Auditory feedback in music performance: The role of melodic structure and musical skill. *Journal of Experimental Psychology: Human Perception & Performance*, **31**, 1331-1345.

Repp, B. H. (1999). Effects of auditory feedback deprivation on expressive piano performance. *Music Perception*, **16**, 409-438.

Roberts, L. D., & Gregory, A. H. (1973). Ear differences and delayed auditory feedback: Effect on a simple verbal repetition task and a nonverbal tapping test. *Journal of Experimental Psychology*, **101**, 269-272.

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

Rosenbaum, D. A., Kenny, S. B., & Derr, M. A. (1983). Hierarchical control of rapid movement sequences. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 86-102.

Rumelhart, D. E., & Norman, D. A. (1982). Simulating a skilled typist: A study of skilled cognitive–motor performance. *Cognitive Science*, **6**, 1-36.

Schmidt, R. A., & Lee, T. D. (1999). *Motor control and learning: A behavioral emphasis* (3rd ed.). Champaign, IL: Human Kinetics.

Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial-ordering mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 295-342). Hillsdale, NJ: Erlbaum.

Sloboda, J. A., Clarke, E. F., Parncutt, R., & Raekallio, M. (1998). Determinants of finger choice in piano sight-reading. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 185-203.

Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition*, **73**, 205-246.

Smith, N. A., & Schmuckler, M. A. (2004). The perception of tonal structure through the differentiation and organization of pitches. *Journal of Experimental Psychology: Human Perception & Performance*, **30**, 268-286.

Stevens, L. T. (1886). On the time sense. *Mind*, **11**, 393-404.

Vousden, J. I., Brown, G. D. A., & Harley, T. A. (2000). Serial control of phonology in speech production: A hierarchical model. *Cognitive Psychology*, **41**, 101-175.

Wing, A. M. (1977). Perturbations of auditory feedback delay and the timing of movement. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 175-186.

Wing, A. M., & Kristofferson, A. B. (1973). The timing of interresponse intervals. *Perception & Psychophysics*, **13**, 455-460.

Yates, A. J. (1963). Delayed auditory feedback. *Psychological Bulletin*, **60**, 213-232.

NOTES

1. We will use the terms *auditory feedback* and *feedback* interchangeably. Although the term *auditory feedback* suggests an association with feedback control theories (see Howell, 2004a), we do not intend to invoke this connotation and, instead, will retain the term for reasons of its connection with past research.

2. Mates and Aschersleben (2000) presented future events in auditory feedback, but on a much shorter time scale than those examined here.

3. Pfordresher (2003) used the term *period shift*, rather than *serial shift*, to emphasize a distinction between period and phase shifts of auditory feedback. Because that distinction is not relevant here, we will use the term *serial shift*.

4. The participants from the Ohio State University community used the Roland RD-600 model in all the experiments, whereas the students from the McGill community (Experiments 1 and 2) and the students from the University of Texas at San Antonio community (Experiment 3) used the Roland RD-700. These keyboard models do not differ in any noticeable way with respect to the sounds used in the experiment or the feel of the keyboard.

5. In general, movement gradients did not differ reliably as a function of the absolute distance between the current event and the auditory feedback event. One exception was found in Experiment 2, for which movement gradients were somewhat flatter (indicating less distinctiveness) for feedback from a distance of 2 (metrically similar events) than for that from other distances.

6. Chance estimates for serial-ordering errors (Y) across distances of 1–3 were computed using the following formula, which assumes equal likelihood of anticipatory and perseveratory errors by chance:

$$Y_x = (n - x) \Big/ \sum_{x=1}^{s} (n - x),$$

where n is the total number of sequence events (12 in the present context), x is the distance between events, and s is the maximal distance analyzed (3 in the present context).

(Manuscript received June 16, 2004;
revision accepted for publication May 12, 2005.)