ORIGINAL ARTICLE

"Deafness" effects in detecting alterations to auditory feedback during sequence production

Peter Q. Pfordresher

Abstract Past research has shown that when discrete responses are associated with a perceptual goal, performers may have difficulty detecting stimuli that are commensurate with that goal. Three experiments are reported here that test whether such effects extend to sequence production. In Experiment 1, participants performed 8-note melodies repeatedly, and on each trial a single tone could be altered with respect to its pitch and/or synchrony with actions. Results suggested a selective deficit of detection when feedback pitch was unchanged and the event was slightly delayed. Experiment 2 showed that this "deafness" to feedback is limited to rhythmic motor tasks that require sequencing, in that similar effects did not emerge when participants produced pitch sequences by tapping a single key repeatedly. A third experiment demonstrated similar results to Experiment 1 when the mapping of keys to pitches on the keyboard was reversed. Taken together, results suggest a selective deafness to response-congruent delayed feedback, consistent with the idea that performers suppress previously planned events during production.

Introduction

When people perform complex auditory sequences such as melodies or spoken sentences, the sequencing of motor movements runs concurrently with the processing of perceptual feedback from these actions, with auditory feedback (the goal of actions) playing a particularly important

role. Past research concerning the effects of altered auditory feedback suggests that performers are particularly sensitive to synchronization between actions and auditory feedback, with delays of auditory feedback causing considerable disruption of production (primarily to the timing of production; for a review see Pfordresher, 2006). Alterations to the pitch content of feedback can also lead to disruption, suggesting that performers are sensitive to deviations of feedback content from the planned outcomes of actions. Although a good deal of research has addressed the way in which alterations to auditory feedback can influence motor planning and production, little is known about how motor planning influences how individuals perceive the timing and pitch content of auditory feedback. I addressed this issue in the experiments reported here, using a paradigm in which a single pitch event may be displaced with respect to the pitch and/or timing that the performer would expect based on a performance plan. Past research has used such alterations to better understand the effects of selected perturbation on performance (Furuya & Soechting, 2010) and neural responses to altered events (Maidhof, Vavatzanidis, Prinz, Rieger, & Koelsch, 2010).

The theoretical motivation for this research stems from recent claims about the way in which perception and production interact. In contrast to claims that action planning may be dependent on the presence of appropriate feedback (e.g., Chase, 1965; Guenther, 1995), or that planning may function in a feed-forward manner with minimal influence of feedback (e.g., Howell, 2001; Howell, Powell, & Khan, 1983; Lashley, 1951), some recent proposals have taken an intermediate stance in suggesting that perception and action may share cognitive resources. By such accounts, perception and production are not dependent on each other but can interact based on the need to sustain a single representation that serves both domains (e.g., Hommel, 2009;

P. Q. Pfordresher (✉)
Department of Psychology, University at Buffalo,
State University of New York, 355 Park Hall,
Buffalo, NY 14260, USA
e-mail: pqp@buffalo.edu

Hommel, Müsseler, Aschersleben, & Prinz, 2001; MacKay, 1987; Shin, Proctor, & Capaldi, 2010). Related computational models of motor control suggest that action planning is guided by a "forward model" that predicts the sensory outcomes of actions, leading to the suppression of perceptual events that are congruent with predictions (e.g., Jordan & Rumelhart, 1992; Kawato, 1999; Wolpert, 1997). With respect to the role of auditory feedback in sequence production, the shared representation approach suggests that disruptive effects of altered feedback occur because feedback events add interfering activation associated with events and/or event timings that conflict with the goals of planning (Pfordresher, 2006). To date, research on perception/action relationships in sequence production has focused exclusively on how perceptual feedback influences production. By contrast, research adopting simpler discrete-response tasks has investigated perception/action relationships in both directions, including the way in which action planning influences perception.

Previous research concerning the effects of action planning on perception suggests that the initiation of an action plan leads to "blindness" (reduction of sensitivity) for visually presented events that are commensurate with the goal of that action (Müsseler & Hommel, 1997a, b; Stoet & Hommel, 1999). Stoet and Hommel (1999), for instance, showed that participants are less able to identify the direction of an arrow when the arrow matched the direction of a previously launched action goal. Such blindness effects extend to both identification and detection (Müsseler & Hommel, 1997b). In the context of feedback dimensions described by Pfordresher (2006), this blindness may be said to relate to the "content" of an action plan (categorical information associated with the goal) as opposed to the timing of an action plan (the time at which one intends to reach the goal, which for a discrete-response task is typically "as soon as possible"). In the context of music performance, then the results of Stoet and Hommel pertain most directly to whether performers can detect feedback pitches associated with actions (here referred to as "deafness" effects given the auditory modality). However, feedback timing may modulate sensitivity to altered pitch, in that the binding of actions with perceptual events may vary over time (Hommel, 2009). How might such temporal binding work in the production of temporally precise sequences, like music?

The task of producing several events sequentially and with high temporal precision leads to constraints on planning beyond those found in discrete-response tasks used in the experiments described above. One critical difference is that sequence production involves the production of several events in a prescribed order. As such, sequence production is thought to involve incremental planning (e.g., Kempen & Hoenkamp, 1987; Palmer & Pfordresher, 2003; Pfordresher, Palmer, & Jungers, 2007; Smith & Wheeldon, 1999;

Wheeldon, Meyer, & Smith, 2002) which involves the activation of multiple sequence events at once, occasionally resulting in serial ordering errors (Lashley, 1951). Of all types of serial ordering errors, perseverations (repetitions of past events) are considered to be indicative of poor planning, and are associated with deficits of sequencing (Dell, Burger, & Svec, 1997). A sequence planning mechanism that has been proposed for the avoidance of perseverations is *post-output suppression* (e.g., MacKay, 1987; Vousden, Brown, & Harley, 2000). Once an event is selected for production, the activation of that event within the performance plan is suppressed, in order to prevent its repetition. Applied to the current study, the phenomenon of post-output suppression leads to a novel hypothesis that was not testable in previous paradigms: that performers may show a selective deficit for detecting feedback events that are delayed but unchanged with respect to pitch. Such events, in principle, would be undergoing post-output suppression, which ought to reduce their detectability. Post-output suppression has been documented previously in speech production, based on a greater tendency for serial ordering errors to be anticipatory rather than perseveratory (Dell et al., 1997; Vousden et al., 2000). By contrast, previous research on serial ordering errors in music suggests that post-output suppression may not occur in music, with similar rates of anticipations and perseverations (Palmer & Pfordresher, 2003; Pfordresher & Palmer, 2006). However, it is possible that the novel paradigm used here may show more subtle suppression effects than has been possible in previous research.

Three experiments are reported here that address how well performers detect alterations to auditory feedback during sequence production. In Experiment 1, participants performed simple melodies on a keyboard, and on half the trials one tone could be shifted up or down in pitch, and/or could be shifted early or late with respect to time. In Experiment 2, participants produced melodies by tapping a single key where each key press generated the next pitch in a melody. Experiment 2 was designed to disentangle the effects of action timing from the planning of action sequences. Experiment 3 addressed the role of matches between sequential key presses and feedback pitch by reversing the relationship between actions and key presses (keys to the right led to lower pitches and vice versa). Participants in all experiments were sampled from Introduction to Psychology classes and were primarily non-musicians. Non-musicians are sensitive to the same feedback pitch alterations as are pianists, though the groups differ with respect to the magnitude of disruption they experience (Pfordresher, 2005, 2012). Moreover, the focus of the present research is not on music performance as an isolated behavior, but rather to understand the effect of action planning on perception of feedback for sequence production in general.

Experiment 1

Method

Participants

Twenty-six individuals participated in exchange for course credit while taking Introduction to Psychology at the University of Texas at San Antonio (UTSA). The mean age of participants was 22 years (range 18–49); 19 participants were male (7 were female); 22 participants reported being right handed (four were left-handed). The sample was predominantly non-musician: 19 participants reported no musical training, and 3 participants reported less than 2 years of training. Mean years of musical training among the remaining 4 participants was 5.6 years (range 3–11), summed across all musical instruments (or voice) reported. Only one participant would be considered a pianist, having 8 years of private lessons and 4.5 weekly hours of current practice. This participant also reported having absolute pitch (no other participants did). The only other participant reporting piano training reported only 3 years of private lessons. All participants reported having normal hearing.

Materials

Each participant learned and then performed from memory one of four melodies that have been used in similar studies (first described in Pfordresher, 2005). All four melodies were designed to be easily learned and performed by non-musicians. Five pitch classes (C to G in the key of C-major; "white keys" on the keyboard) were arranged to form 8-note melodies that varied with respect to contour shape (smooth versus alternating) and starting pitch (C or G). All melodies were designed to be played repeatedly without pausing between repetitions. In lieu of standard notation, melodies were displayed during learning as rows of numbers referring to finger-key combinations, numbers corresponding to the appropriate keys were displayed above the keyboard. There were no changes in hand position, thus finger-key mapping was constant throughout each melody. Participants only viewed notation during the learning phase (**see "Procedure").

Apparatus and conditions

Participants performed on a FATAR CMK 20 MIDI controller keyboard. MIDI output from the keyboard was fed into a Linux-based computer running the program FTAP (Finney, 2001). After processing by FTAP, MIDI output went to a Roland RD-700 keyboard, which functioned as a tone generator. Feedback tones originated from the program 1 ("concert piano") sound patch with all effects (reverberation, chorus) turned off. Participants listened to auditory feedback over Sony MDR-7500 professional headphones, set at a comfortable listening level. The intensity of feedback tones was set to be constant; thus participants could not control feedback intensity through changes in finger velocity.

Participants performed each melody three times in a trial. On half the trials, a single tone was displaced during the second repetition, at the 3rd, 5th or 7th serial position within the melody, whereas no feedback events were altered on the remaining trials. These are referred to as "probe positions." Changes at each probe position could be one of 8 types, resulting from displacements of time (early or late) while keeping pitch normal, displacements of pitch (up or down) while keeping feedback synchronous, or joint displacements (four more conditions). Upwards pitch changes were implemented by having FTAP shift the pitch associated with that serial position up one semitone, whereas pitch was shifted down one semitone from the intended pitch for downward pitch changes. Late asynchronous feedback was implemented by inserting a delay equal to 33 % of the anticipated inter-onset interval (IOI) for that event, based on the average of the two preceding tones (see Pfordresher, 2003 for more details). On trials where feedback was both late and altered with respect to pitch, FTAP both altered the MIDI note number for feedback and delayed its onset. By contrast, early asynchronous feedback was presented 33 % early, based on the anticipated IOI to precede the probe tone and no feedback resulted from the keypress at the probe position.[1]

Crossing the 3 probe positions with 8 change types led to 24 trials with some kind of altered feedback event. Each participant experienced all 24 change trial types for the melody they performed. An additional 24 no-change trials were added for a total of 48 trials per participant. Each participant was randomly assigned to 1 of 4 stimulus melodies and 1 of 2 random orders of trials.

---

[1] Early asynchronies were complicated to implement. For early asynchronies that were associated with the correct (unaltered) pitch, FTAP generated normal auditory feedback for the keypress that preceded the probe position; in addition, that keypress generated a second delayed feedback event that had a MIDI pitch matching the pitch associated with the probe position. Then, when the participant pressed the key at the actual probe position, no feedback would be presented. The delay preceding the probe was set to be 66 % of the predicted IOI. Thus, participants experienced the illusion that the tone they produced occurred early (by 33 % of the IOI preceding the probe position), when in fact the tone was associated with delayed feedback, with altered pitch, from the previous keypress. A similar procedure was followed for early asynchronies combined with an altered pitch, only in such cases the pitch of delayed feedback from the preceding keypress was shifted up or down relative to the pitch intended for the probe position.

Procedure

After participants provided informed consent, the experimenter explained the notation system to participants, and had them play the stimulus melody once to ensure that the participant understood the notation (all did). Participants were then allowed to practice the melody with normal auditory feedback (heard over headphones) until they thought they could perform the melody from memory. After indicating this to the experimenter, the notation was removed, and the participant attempted to play three consecutive error-free renditions of the melody from memory. If they were able to, the experimenter proceeded to the next stage. If not, the participant was allowed more time to practice.

Following the learning stage, participants experienced one practice trial with an altered pitch event. At the beginning of a trial, 4 metronome clicks sounded (using a cowbell timbre from program 126 of the RD-700, corresponding to MIDI note 56). Participants were informed that these clicks set the tempo, and that they should start playing at this tempo when they would expect the 5th click to occur. Participants were instructed to play through the melody with legato articulation (described as keeping one's finger on a given key until they play the next key) until auditory feedback ceased, which signaled the end of the trial. The practice trial featured an altered pitch at probe position 7, which was shifted late in time and upwards in pitch.

After completing a trial, participants made the following reports. First, participants reported whether there was an altered pitch. If the participants did not hear a feedback alteration, the next trial began. However, if the participant did report hearing a feedback alteration, they were asked three more questions. First, participants reported which position (out of 8 possible) in the sequence they thought the alteration occurred. Next, participants reported whether the pitch of the event was shifted up or down (or not altered). Finally, participants were asked whether the timing of the event was synchronized with the keypress, was earlier, or was later. Participants responded to all these items on the practice trial. If a participant did not get one of the responses correct, the experimenter informed him or her of the correct response and then allowed the participant to repeat the practice trial.

Following the practice trial, participants completed the first 24 trials of the experiment. Then participants took a break and completed questionnaires concerning their musical background before completing the next 24 trials.

Data analysis

The primary variables of interest in this study were self-report data collected after the trial: detection of changes to feedback, reported positions of changes (if present), the direction of pitch changes (if present), and the direction of temporal asynchronies (if present). Performance data were also collected via MIDI, comprising timing of key presses (analyzed as inter-onset intervals IOIs), sequencing of pitch events, and keypress velocities. Sequencing data were used to determine the frequency and location of sequencing errors, for the purpose of filtering out problematic trials (see below). Such errors included deletions and additions, both of which shift the position of events subsequent to the error and cause the pitch produced at the probe position to differ from the intended pitch. Errors were analyzed using a suite of programs designed to compare produced sequences to the ideal performance, accounting for the minimum number of changes necessary to account for discrepancies between the performances (Large, 1993; Palmer & van de Sande, 1993, 1995).

Results

Preliminary analyses indicated no significant effect of probe position, and no reliable differences across conditions with respect to the position at which changes were reported to occur, for trials in which the participant detected a change of feedback. Thus, analyses focused on the detection of changes to auditory feedback (i.e., reports of the presence or absence of a change) and labeling responses for the type of change, averaging across probe positions. The ANOVA design used for all tests reported below was a factorial design based on crossing three levels of Timing (feedback timing for probe = early, on time, late) with three levels of Pitch (feedback pitch for probe = shifted up, normal, shifted down).

As described above, trials with addition or deletion errors prior to the probe position were removed. Dropped trials comprised 7 % of all trials (84 dropped trials). This filtering process led to missing data in 4 cells from the matrix based on crossing participants with feedback condition (26 participants × 9 feedback conditions = 234 cells). Missing data were replaced with the imputation procedure described in Keppel and Wickens (2004, pp. 395–399). Briefly, for a given cell reflecting the intersection of a participant and condition, this procedure generates an estimate based on mean performance across all other participants for that condition, as well as mean performance across all other conditions for that participant.

Participants' detection responses (whether a change was present or absent) were converted to accuracy scores; means across conditions are shown in Fig. 1. The ANOVA yielded a significant main effect of Pitch, $F(2, 50) = 29.43$, $p < .01$, $\eta_p^2 = .54$, with higher accuracy for conditions with some kind of change (M accuracy for upwards change = 93 %, SE = 2 %, for downwards change $M = 91$ %, SE = 2 %) than for trials with no change ($M = 78$ %, SE = 3 %). In addition, there was a

Fig. 1 Effects of feedback pitch and time alterations on the accuracy of detecting the presence or absence of alterations to auditory feedback in Experiment 1. *Standard error bars* reflect one standard error of the mean
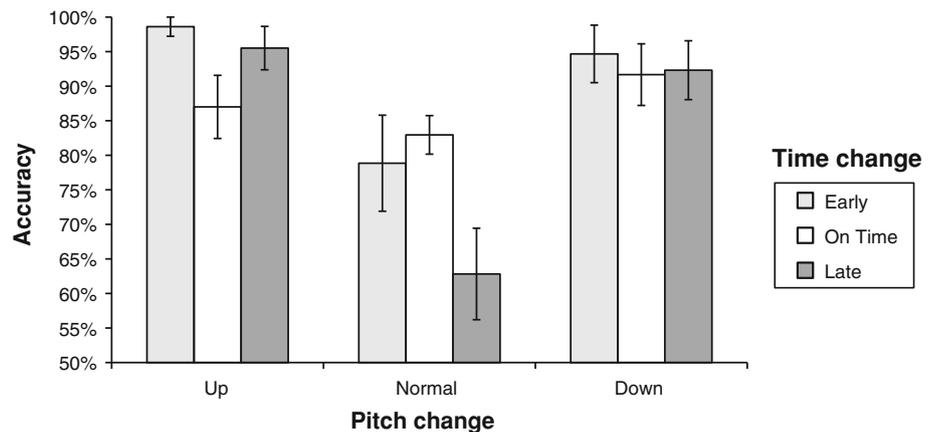
significant Pitch × Timing interaction, $F(4, 100) = 3.05$, $p < .05$, $\eta_\mathrm{p}^2 = .11$. The interaction was driven by the sharp decrease in accuracy for feedback alterations in which the probe tone was delayed but was not altered in pitch. The main effect of Timing was not significant ($p = .20$, $\eta_\mathrm{p}^2 = .06$). Note that the responses shown in Fig. 1 only coded accuracy in detecting the presence or absence of a change, and did not take into account how participants labeled the type of change (which is discussed later).

An important issue in past research within the domain of music cognition has concerned whether either of the primary musical dimensions—typically considered to be pitch and time—is more dominant in the context of a particular task (e.g., Ellis & Jones, 2009; Hébert & Peretz, 1997; Krumhansl, 2000; Prince, Schmuckler, & Thompson, 2009; Prince, Thompson, & Schmuckler, 2009). Recent evidence suggests a dominant role for pitch in retrieval of musical sequences (Prince & Pfordresher, 2012). To test whether changes to pitch in the present study were more or less salient than changes to timing, we compared accuracy for trials with only a pitch change (when timing was synchronous) versus those trials with only a time change (where pitch was normal). In keeping with the pitch dominance hypothesis, accuracy was significantly higher for trials with a changed pitch [$M = 89\%$, SE = 4 %, for time changes, $M = 71\%$, SE = 5 %, $t(25) = 3.67$, $p < .01$, $r^2 = .35$].

Figure 2 shows participants' labeling of changes to pitch (a) or time (b), which were assumed to be "no change" if a participant reported that no alteration was present (for a breakdown of responses into response categories see Table 1 in Appendix). Reported upwards pitch changes were assigned a value of 1 and downwards changes a value of −1. As can be seen in Fig. 2a, there was an overall bias for downwards responses ($M = -.16$). Labeling of pitch changes (shown in Fig. 2a) yielded significant main effects of Pitch, $F(2, 50) = 6.91$, $p < .01$, $\eta_\mathrm{p}^2 = .22$, Timing, $F(2, 50) = 3.35$, $p < .05$, $\eta_\mathrm{p}^2 = .12$, and a Pitch × Timing interaction, $F(4, 100) = 4.04$, $p < .01$, $\eta_\mathrm{p}^2 = .14$. The main

effect of Pitch reflects the fact that participants, appropriately, were more likely to report a downwards pitch change when feedback was in fact shifted down ($M = -.38$, SE = .06) than when pitches were shifted upwards ($M = -.05$, SE = .07) or there was no change of pitch ($M = -.07$, SE = .03). Of greater interest was the fact that the timing of feedback influenced reports of changes to pitch. When feedback onsets were delayed (presented late), there was a greater tendency to label feedback pitch as shifted downwards ($M = -.24$, SE = .06) than when feedback was presented early, which led to responses closer to neutrality ($M = -.08$, SE = .06). Interestingly, pitch responses were more accurate (closer to zero) for early feedback events than on time events, which like delayed events elicited a bias to label feedback as shifted down in pitch ($M = -.19$, SE = .05). The interaction was analyzed via simple effects of pitch within each feedback timing condition. Whereas the main effect of pitch was preserved for on time and delayed feedback conditions, the main effect of pitch was not significant for early shifts. Thus, participants found it difficult to identify the relationship of feedback pitch to the intended pitch, when the feedback pitch occurred slightly before the intended pitch.

Labeling responses for changes to timing, shown in Fig. 2b, led to a simpler pattern of results than we found for labeling of changes to pitch. The ANOVA yielded only a significant main effect of Timing, $F(2, 50) = 22.73$, $p < .01$, $\eta_\mathrm{p}^2 = .48$, whereas the main effect of Pitch and the interaction were both negligible ($F < 1$ in both cases, main effect of pitch $\eta_\mathrm{p}^2 < .01$, interaction $\eta_\mathrm{p}^2 = .02$). Thus, feedback pitch did not bias labeling of time in the way that feedback timing biased labeling of pitch. The main effect of time simply reflected the fact that participants correctly labeled early asynchronies as being early ($M = .22$, SE = .07, note that early responses were assigned a label of +1), late responses were given for delayed asynchronies ($M = -.43$, SE = .06), and responses with no time change were typically labeled as such ($M = .02$, SE = .04).

Fig. 2 Effects of feedback pitch and time alterations on labeling responses for pitch (**a**) and timing (**b**) of probe events in Experiment 1. *Standard error bars* reflect one standard error of the mean
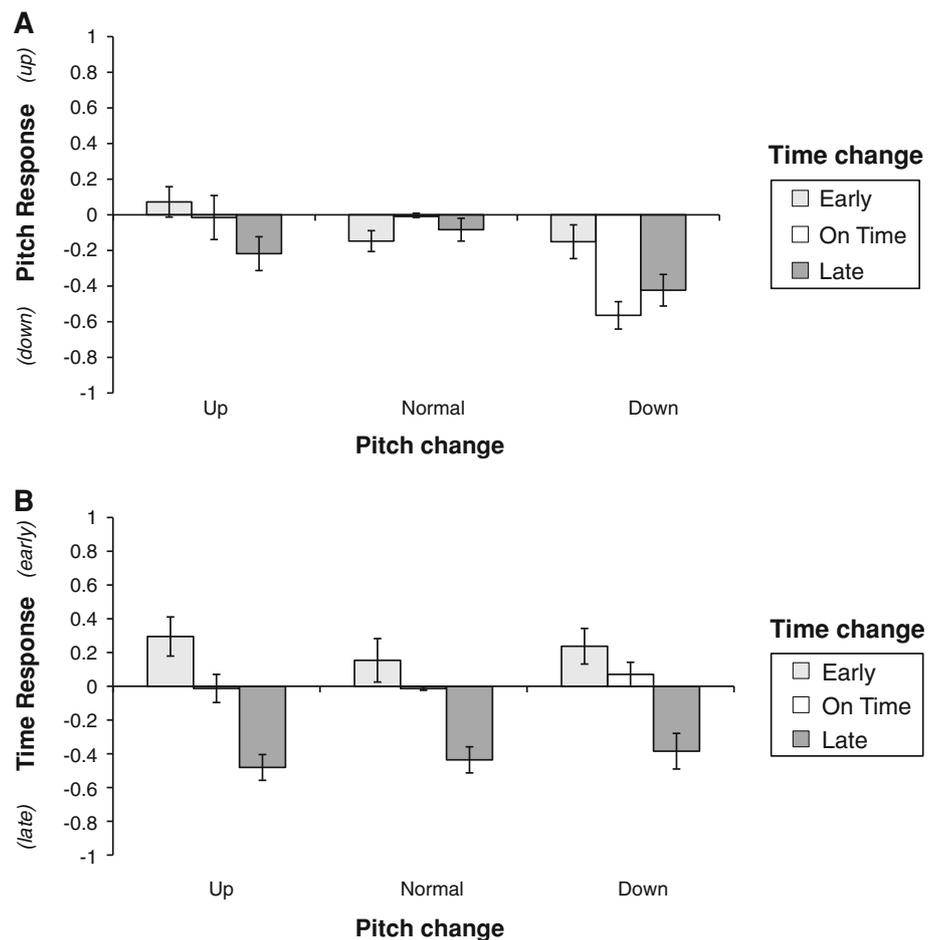
Discussion

Results from Experiment 1 revealed that participants' ability to detect and label feedback events during sequence production varies as a function of the type of alteration. The most critical result, theoretically, is the fact that events that appeared late but that were unaltered with respect to pitch were detected with lowest accuracy. This result is consistent with findings of "blindness" for perceptual events that are congruent with a planned response. Similar to the result of Stoet and Hommel (1999), detection was less accurate when pitch remained unaltered (and was thus consistent with the most recent action goal), then when pitch was altered. Beyond this, results suggested post-output suppression: detection was particularly disadvantaged when the unaltered probe pitch was slightly delayed. In such cases, the performer in theory actively suppresses activation associated with that pitch event, thus inhibiting its detectability beyond what was found when feedback was on time or early. Labeling of pitch changes further suggests a vulnerability of pitch labeling to feedback timing, whereas the converse was not found in temporal labeling responses. Though this finding seems to run

against previous results suggesting a dominant role of pitch over time in sequence production (Prince & Pfordresher, 2012), accuracy data nonetheless suggest that pitch changes were overall more salient than feedback asynchronies.

Results of Experiment 1 thus suggest that the task of planning and producing action sequences leads to difficulty in processing pitch contents of feedback events. This conclusion is limited, however, by the fact that all trials from Experiment 1 involved sequence production. Furthermore, because the task of Experiment 1 required participants to sequence actions correctly and with high temporal precision, it is unclear how much the results reflect sequencing versus motor timing. Therefore, Experiment 2 removed sequencing constraints from the motor task by having participants tap isochronously, with each subsequent tap producing a new tone of a melody.

Experiment 2

Experiment 2 was identical to Experiment 1 with the exception that participants tapped the same key repeatedly on each trial instead of producing finger sequences. Each

tap on a trial generated a new pitch in a melody, using the same set as in Experiment 1. Thus, feedback events in Experiment 2 were identical to those in Experiment 1, even though the motor task differed. The same feedback alterations were presented as in Experiment 1. Alterations to pitch in Experiment 2 were judged relative to the pitch that ought to occur in the melody, rather than the pitch associated with the participant's keypress.

Method

Participants

Twenty-two individuals participated in exchange for course credit while taking Introduction to Psychology at the University at Buffalo (UB). Due to an experimenter oversight, general demographic information (age, sex, handedness) was not recorded for 6 participants, though musical background data were recorded for all 22 participants. The mean age of participants having demographic data was 19 years (range 18–23); 7 of these participants were male (9 were female) and all participants reported being right handed. In comparison with Experiment 1, more participants from Experiment 2 reported musical training, though not necessarily piano training. Nine participants reported 2 years of training or fewer (6 reported no training whatsoever). Mean years of musical training among the remaining 13 participants was 8.5 years (range 3–21), summed across all musical instruments (or voice) reported. Across all participants, the mean years of reported music lessons in Experiment 2 ($M = 5.2$) was higher than in Experiment 1 ($M = 2.9$), $t(46) = 3.11$, $p < .01$, $r^2 = .17$. Thee participants in Experiment 2 reported having more than 6 years of private piano lessons (8, 12, and 21 years reported), and were thus considered pianists (recall, however, that Experiment 2 did not involve production of melodies on a keyboard). One participant (the pianist reporting 8 years of training) reported having

absolute pitch. All participants reported having normal hearing.

Apparatus and conditions

The same experimental setup was used as in Experiment 1. A minor change of apparatus in Experiment 2 (which was carried out at a different institution) is that performers used an M-AUDIO Keystation 49e controller for tapping responses (a very similar device to the FATAR keyboard used in Experiment 1). In Experiment 2, FTAP generated pitches from a vector of 8 pitch events that repeated through a trial, such that each key press generated the next pitch in the sequence (cycling back to the first position after the sequence was finished).

Procedure

The procedure was identical to Experiment 1; with the exception that participants were told to judge feedback pitches relative to the next pitch they would expect to hear in the melody, which they learned by tapping it and hearing pitch events in the learning phase of the experiment.

Results

As in Experiment 1, no effects of probe position, or differences in reports of probe position, were evident and responses were analyzed using a 3 (Timing) × 3 (Pitch) within-participants ANOVA. Because sequencing errors were not a concern for Experiment 2, we were able to use all trials in analyses and no accommodation for missing data was necessary.

Figure 3 shows mean accuracy in detecting the presence or absence of a feedback alteration as a function of feedback pitch and timing, displayed as in Fig. 1. There was a main effect of Timing, $F(2, 42) = 14.71$, $p < .01$, $\eta_p^2 = .41$, whereas the main effect of Pitch and the

Fig. 3 Effects of feedback pitch and time alterations on the accuracy of detecting the presence or absence of alterations to auditory feedback in Experiment 2. *Error bars* reflect one between-participants standard error of the mean
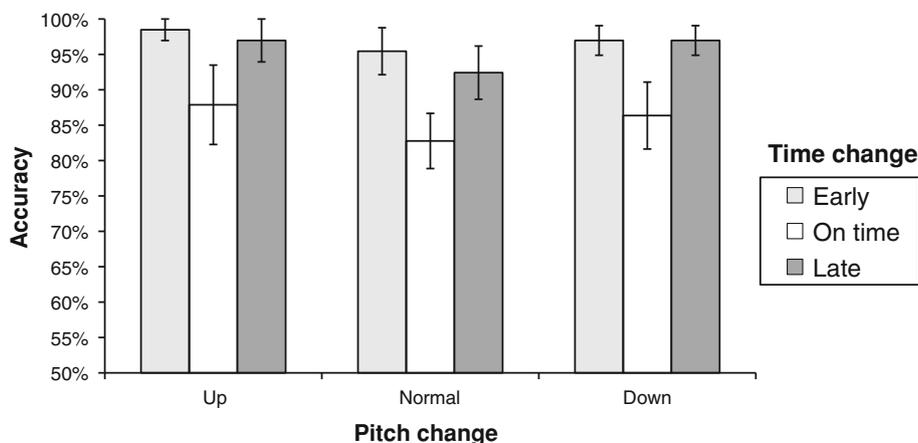
interaction were non-significant ($p > .10$ in each case, main effect of pitch $\eta_p^2 = .09$, interaction $\eta_p^2 < .01$). Participants were more accurate at detecting asynchronous feedback that was either early ($M = 95\%$, SE = 2 %) or late ($M = 97\%$, SE = 1 %) than they were at detecting synchronous feedback ($M = 86\%$, SE = 3 %). By contrast, participants' processing of feedback pitch was similarly accurate across all conditions. In addition, the test for pitch dominance in accuracy data, carried out as in Experiment 1, went in the opposite direction of the Experiment 1 data, with higher accuracy for asynchronous feedback ($M = 93\%$, SE = 3 %) than for changes to pitch ($M = 87\%$ SE = 5 %), though this difference was not significant ($p = .08$).

Figure 4 shows labeling responses based on feedback pitch (a) and timing (b) (for a breakdown of responses into response categories see Table 2 in Appendix). Similar to accuracy data, these results are straightforward and do not show the same kind of cross-dimensional influence that was found for pitch label responses in Experiment 1. The ANOVA on pitch label responses yielded a significant main effect of Pitch, $F(2, 42) = 13.33$, $p < .05$, $\eta_p^2 = .39$, but no main effect of Timing ($p = .09$, $\eta_p^2 = .11$), and no interaction ($p = .68$, $\eta_p^2 = .03$). Participants labeled

pitches appropriately and were not influenced by feedback timing. Likewise, the ANOVA on time label responses yielded only a main effect of Timing, $F(2, 42) = 26.62$, $p < .01$, $\eta_p^2 = .56$, but no main effect of Pitch ($p = .34$, $\eta_p^2 = .05$), and no interaction ($p = .09$, $\eta_p \eta_p^2 = .09$).

Pooled results across Experiments 1 and 2

Results of Experiment 2 thus suggest that the need to time movements with rhythmic regularity does not lead to the same "deafness" effects as found in Experiment 1. A possible concern with this interpretation is that differences across experiments may reflect differences in the difficulty of the sequencing versus tapping tasks, which could influence attentional resources one can devote to auditory feedback. To address these issues, I conducted analyses on pooled data across experiments using a 2 (Experiment) × 3 (Timing) × 3 (Pitch) mixed-model ANOVA. Because results for labeling of time did not differ across experiments, analyses were conducted on detection accuracy and pitch labeling responses, focusing on main effects and interactions with the factor Experiment (effects not associated with Experiment are omitted for sake of brevity).

Fig. 4 Effects of feedback pitch and time alterations on labeling responses for pitch (a) and timing (b) of probe events in Experiment 2. *Error bars* reflect one between-participants standard error of the mean
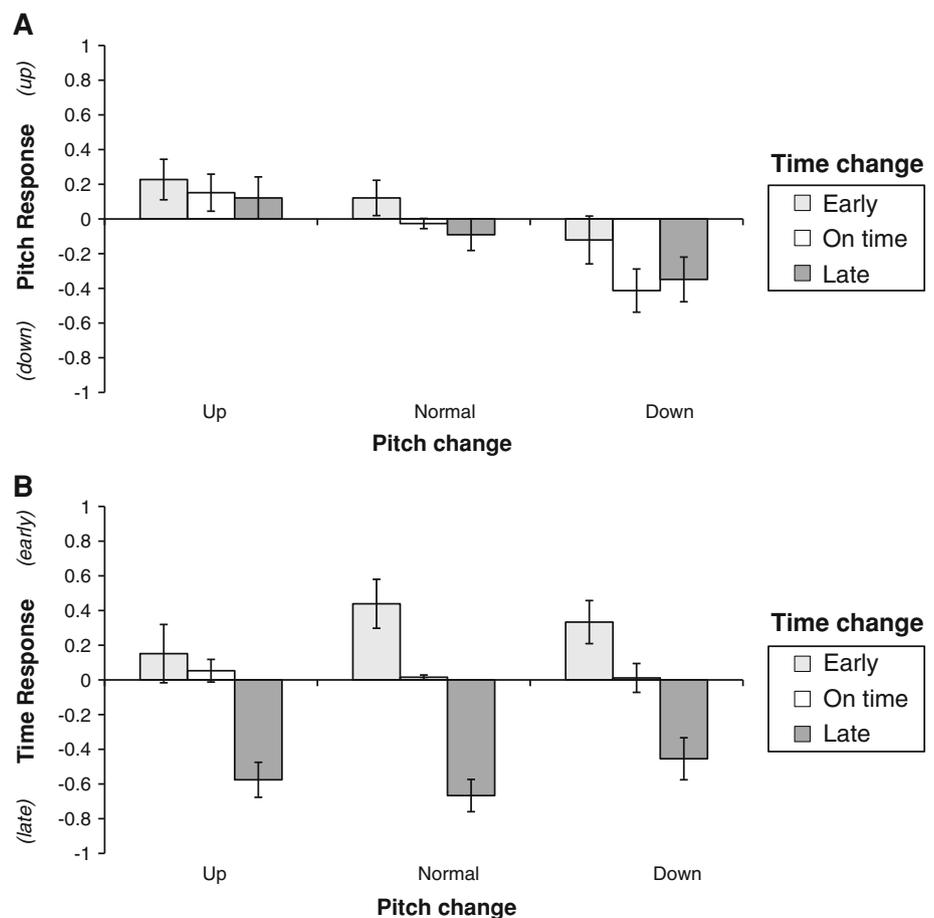
The ANOVA on detection accuracy yielded a significant Experiment × Timing interaction, $F(2, 92) = 4.78$, $p < .05$, $\eta_p^2 = .09$, and a significant Experiment × Pitch interaction, $F(2, 92) = 10.99$, $p < .01$, $\eta_p^2 = .19$, but no main effect of Experiment ($p = .10$, $\eta_p^2 = .06$) and no 3-way interaction ($p = .11$, $\eta_p^2 = .04$). The two significant interactions are plotted in Fig. 5. For the dimension of timing (Fig. 5a), detection of asynchronies was less accurate in Experiment 1 than in Experiment 2, whereas participants in both experiments were similarly able to detect synchronous feedback. With respect to detection of pitch changes, responses to unchanged events were less accurate in Experiment 1 than Experiment 2, whereas participants in both experiments were similarly accurate at detecting changes to pitch.

The role of musical training

As noted in the "participants" section above, the sample from Experiment 2 contained more musically trained individuals than in Experiment 1. To ensure that "deafness" effects apparent in Experiment 1 (but not Experiment 2) for pitch were not due to differences in the samples, I re-ran the ANOVA reported above on only those participants reporting 2 years of training or fewer; this criterion provided a reasonable number of participants in each sample ($N = 22$ from Experiment 1, $N = 9$ from Experiment 2), while keeping reported years of training to a minimal level. Results mirrored those from the ANOVA on all participants, with both interactions retaining significance and reflecting the same trends. To quantify similarity, I correlated means from the 3-way interaction across

both ANOVAs. This correlation suggested a considerable degree of similarity, $r(16) = .98$, $p < .01$.

I also conducted an ANOVA with the same design on labeling of pitch responses. Results within each experiment suggest that feedback timing biased pitch responses for the sequencing task of Experiment 1, but did not have the same effect for the tapping task in Experiment 2. However, these differences were not robust enough to lead to a significant Experiment × Timing interaction in the pooled analysis.

A potentially important difference across experiments stemmed from differing motor demands of the tasks and their influence on timing. Specifically, non-pianists are highly variable with respect to timing when attempting to reproduce musical sequences on a keyboard, in comparison to performance of such participants on the much simpler tapping task (cf. Pfordresher, 2005; with Pfordresher & Dalla-Bella, 2011). Likewise, in the current study, timing variability (measured using coefficient of variation, $CV = SD/M$ for inter-onset intervals in a trial) was significantly higher in Experiment 1 ($M = .21$, $SE = .03$) than in Experiment 2 ($M = .04$, $SE < .01$), $t(44) = 5.05$, $p < .01$, $r^2 = .37$.

This difference in task constraints may have influenced asynchrony detection in Experiments 1 and 2. Though asynchronies were set up to be ±33 % of the anticipated inter-onset interval, it is impossible to control feedback timing in this way with complete accuracy, and timing variability directly influences the precision of this manipulation. Thus, a follow-up analysis was run on trials from each experiment in which the observed asynchrony reflected the intended asynchrony to an optimal degree. Feedback asynchronies, for trials in which feedback was
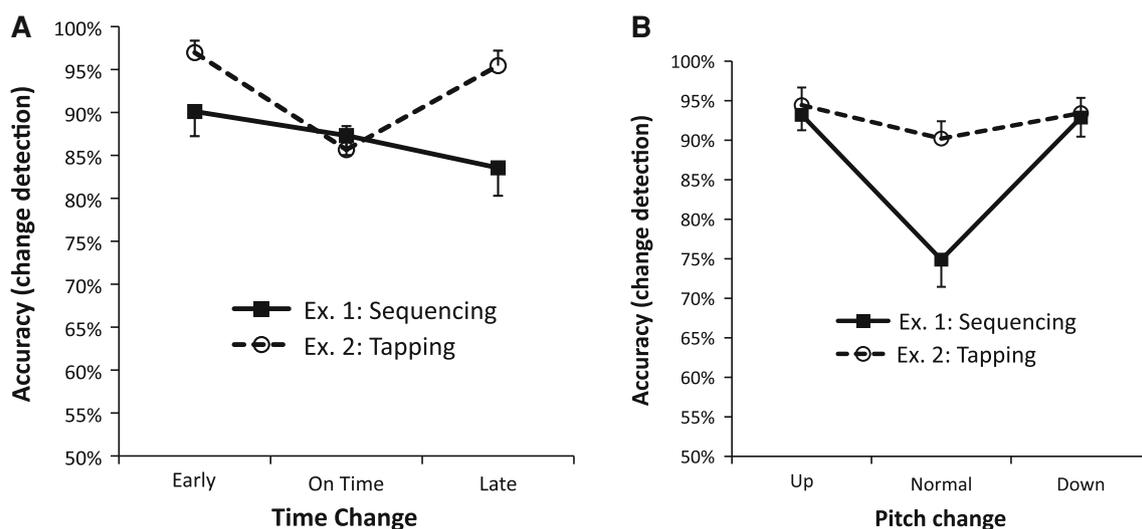
Fig. 5 Significant two-way interactions between the factors Experiment and Timing (**a**), and the factors Experiment and Pitch (**b**) in the ANOVA comparing Experiments 1 and 2. *Error bars* reflect one between-participants standard error of the mean

asynchronous, were categorized into bins of 10 % based on relative feedback asynchronies (magnitude of delay/length of inter-onset interval). For both experiments, the modal relative asynchrony for late feedback events was between 20 and 30 % of the inter-onset interval, slightly earlier than the prescribed delay (33 %). Accuracy for these trials from Experiment 1 was lower than for comparable trials in Experiment 2, as in the mean across all trials reported above (M for Experiment $1 = 75$ %, SE = .95, for Experiment 2 $M = .95$, SE = .03), $t(44) = -2.09$, $p < .05$, $r^2 = .09$. For early feedback events, modal performance in each experiment was between 30 and 40 % early relative to the inter-onset interval, as prescribed, and detection accuracy in Experiment 1 was similar to that of Experiment 2 in these trials (Experiment 1 $M = .99$ SE = .01, Experiment 2 $M = .98$, SE = .01). Thus, the influence of the sequencing task on asynchrony detection was asymmetric, with the decrement being valid for detection of late events more so than early events.

Discussion

Results from Experiment 2 revealed that the task of sequencing actions contributes to feedback processing beyond any effects associated with the motor timing component of the task. Differences across experiments came in the form of improved detection of pitch changes during the tapping task, as opposed to the sequencing task of Experiment 1. It is tempting to attribute these differences to overall task difficulty, given the inherently greater demands of sequence planning, even with simple 8-event repeating sequences such as those used here. However, the absence of any main effects associated with experiment, particularly on overall accuracy, argues against such an interpretation. Moreover, results from Experiment 2 did not demonstrate the selective deafness to delayed feedback events with an unchanged pitch. This is a critical point in that late asynchronies are predicted to be more strongly affected by the sequencing task than are early asynchronies, given the assumption of post-output suppression in sequence production.

Thus, taken together, the results of Experiments 1 and 2 suggest that planning motor sequences influences the ability to detect certain alterations to auditory feedback, beyond the need to maintain temporal regularity. However, one difference across these experiments has to do with the mapping of actions to sounds. Whereas in Experiment 1, the spatial location of actions was commensurate with pitch height, action selection was independent of pitch in Experiment 2. A third experiment was conducted to test whether "deafness" effects associated with sequencing can be found when the mapping of pitch height to the spatial location of keys differs from the standard mapping of actions to sound on a keyboard.

Experiment 3

In Experiment 3, pitch classes on the keyboard were reversed with respect to pitch height. Thus, the spatial locations on the keyboard were negatively correlated with the standard key-to-pitch mapping (where left = low pitch and right = high pitch). Past research has shown that musically untrained individuals are usually sensitive to the standard keyboard pitch mapping (Lidji, Kolinsky, Lochy, & Morais, 2007; Rusconi, Kwan, Giordano, Umilta, & Butterworth, 2006). We chose this kind of mapping, as opposed to a more arbitrary mapping (e.g., scrambling pitch), for two reasons. First, we wanted the feedback sequence to be perceptually coherent as an auditory object, so that participants could judge alterations to feedback at an acceptable level of accuracy. Second, certain alterations of perception/action mapping can be disruptive to action planning (see Pfordresher, 2006, for a review), and in this experiment it was important that action planning be performed fluently.

At issue was the flexibility of perception/action mapping in participants' ability to process auditory feedback events. One possibility is that these associations are firmly entrenched, and that any alteration of the standard pitch-feedback mapping reduces deafness effects seen in Experiment 1. If so, results from Experiment 3 should mirror those from Experiment 2. An alternative possibility is that these associations are flexible to an extent, and that non-disruptive alterations (such as the reversed mapping used here) can lead to a re-mapping of perception/action associations and thus preserve deafness effects seen in Experiment 1.

Method

Participants

Twenty-five individuals participated in exchange for course credit while taking Introduction to Psychology at UB. However, for reasons explained below, the data from three participants were excluded from the sample after analysis of error data. The mean age of the remaining 22 participants was 19 years (range 18–21); 16 participants were male (4 were female and 2 declined to report); 20 participants reported being right handed (1 was left-handed and 1 declined to report). Thirteen participants reported 2 years of training or fewer (10 reported no training whatsoever). The mean years of reported music lessons (summed across all instruments reported) among the remaining 9 participants was 8.9 years (range 3–26). Of those reporting musical training, only one could be considered a pianist (at an intermediate level), having 5 years of private lessons, but this person did not report being

currently active. All participants reported having normal hearing. One participant reported having absolute pitch. However, this participant also reported no musical training; thus, it is possible that the participant did not read the questionnaire carefully, which defined absolute pitch ability in traditional terms (as being able to retrieve a note name upon hearing a pitch).

Apparatus and conditions

The same apparatus as in Experiment 1 was used. As in Experiment 2, FTAP was used to generate a fixed sequence of pitches, with the mapping of pitches reversed so that pressing the key associated with "C" led to the "G" pitch, pressing the "G" led to "C", Pressing the "D" key led to "F", etc.

Procedure

The procedure was identical to Experiment 1. During the learning phase, participants heard auditory feedback with the standard mapping of pitch height to key (left = low pitch, right = high pitch), in order to reinforce this mapping and enhance any effect of its reversal. This mapping was reversed for the practice trial and all experimental trials.

Results

As in Experiment 1, no effects of probe position were evident and responses were analyzed using a 3 (Timing) × 3 (Pitch) within-participants ANOVA. Also as in Experiment 1, sequencing errors during trials were analyzed and any trials with addition or deletion errors before the probe position were removed. Dropped trials comprised 9 % of all trials in total (165 dropped trials). As mentioned

before, three participants were missing enough trials that they lacked any observations in more than one condition. Such loss of data makes the process of imputing missing values questionable (given that the process assumes you have a good mean estimate across conditions for each participant), and suggests poor learning among these participants (most participants performed with high accuracy). After removing these three participants, only one cell required imputation, following the procedure used in Experiment 1. Surprisingly, two of the three participants with enough missing data to warrent exclusion were piano players, which may reflect the greater sensitivity to pitch mapping found in pianists (Pfordresher, 2012). Among those participants who were retained in the sample, the mean percent of errors per trial was equivalent to Experiment 1 (4 %) after rounding to the nearest percent. Thus, as anticipated, the reversed feedback mapping in general does not seem to be disruptive. Most important, any differences between this Experiment and Experiment 1, for the sample used here, should not simply reflect disruptive effects of reversed feedback on performance.

Figure 6 shows mean accuracy for detection of a change as a function of feedback pitch and timing. Results were highly similar to those found in Experiment 1 (see Fig. 1), reflected in a strong positive correlation between the 9 treatment condition means in Experiment 1 and those shown here, $r(7) = .88$, $p < .01$. As in Experiment 1, there was a main effect of feedback pitch, $F(2, 42) = 11.26$, $p < .01$, $\eta_p^2 = .35$, and a significant pitch × timing interaction, $F(4, 84) = 6.03$, $p < .01$, $\eta_p^2 = .22$, but the main effect of timing was not significant ($p = .11$, $\eta_p^2 = .10$).

Figure 7a shows mean pitch labeling responses from Experiment 3 (for a breakdown of responses into response categories see Table 3 in Appendix). Again, means across conditions were highly correlated with those from Experiment 1, $r(7) = .90$, $p < .01$. The ANOVA on pitch label

Fig. 6 Effects of feedback pitch and time alterations on the accuracy of detecting the presence or absence of alterations to auditory feedback in Experiment 3. *Error bars* reflect one between-participants standard error of the mean
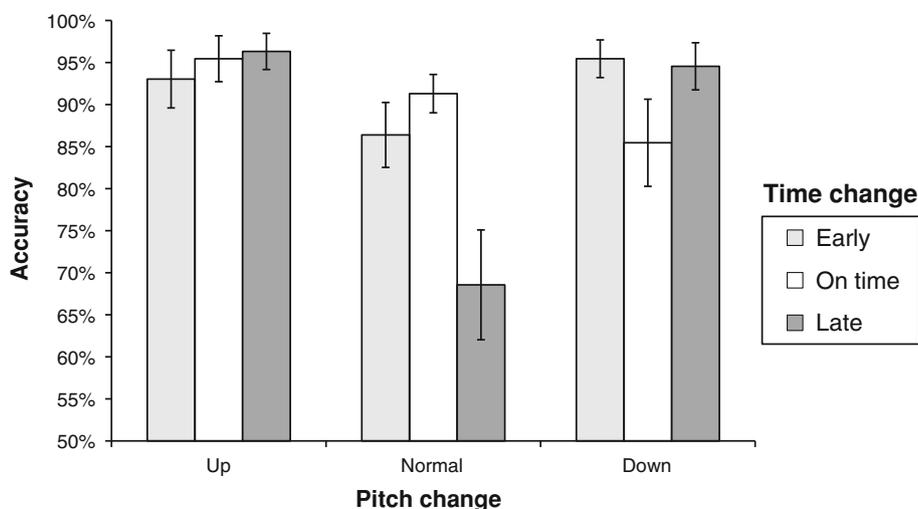
responses yielded a significant main effect of Pitch, $F(2, 42) = 22.40$, $p < .01$, $\eta_p^2 = .52$, and a Pitch × Timing interaction, $F(4, 84) = 3.70$, $p < .01$, $\eta_p^2 = .15$. Unlike Experiment 1, there was no main effect of Timing ($p = .30$, $\eta_p^2 = .06$). However, means reflected a similar trend, with reduced ability to label pitch when feedback was presented early then when it was presented on time or late.

Labeling responses for feedback timing, not surprisingly, mirrored results found for Experiments 1 and 2 and cell means were highly similar to those from Experiment 1, $r(7) = .99$, $p < .01$. As in other experiments, the ANOVA yielded a main effect of Timing, $F(2, 42) = 25.56$, $p < .01$, $\eta_p^2 = .58$, with no main effect of Pitch ($p = .79$, $\eta_p^2 = .01$), and no interaction ($p = .23$, $\eta_p^2 = .06$).

Effects of musical training

Because results from Experiment 3 were highly similar to those from Experiment 1, we pooled the data from these experiments to address a remaining issue: whether feedback deafness varies with musical training. As in follow-up analyses comparing Experiments 1 and 2, we categorized every participant who reported more than 2 years of private musical training as a "musician" ($N = 13$) and the remainder as "non-musicians" ($N = 35$). This liberal criterion for categorization as musician was applied to maximize statistical power in comparisons across groups. We submitted these data to a 3-way mixed-model ANOVA with the factors Group, Timing and Pitch. The main effect of group and interactions with this factor are of primary importance.

The ANOVA yielded a significant Group × Pitch interaction, $F(2, 92) = 5.03$, $p < .01$, $\eta_p^2 = .10$, but no other effects associated with Group (main effect of Group: $p = .13$, $\eta_p^2 = .06$, Group × Timing: $p = .25$, $\eta_p^2 = .03$, Group × Timing × Pitch: $p = .28$, $\eta_p^2 = .03$). The significant 2-way interaction is plotted in Fig. 8. As can be seen, deafness for normal feedback pitch was only found for non-musicians. A series of planned contrasts between groups was also run for each of the nine experimental conditions. The only contrast that was significant was for feedback events that were delayed but not changed in pitch, $t(46) = 5.17$, $p = .03$, $r^2 = .37$. The pronounced feedback deafness found for this condition was again only present for non-musicians ($M = 59\%$, $SE = 5\%$) and not for musicians ($M = 82\%$, $SE = 7\%$).

Fig. 7 Effects of feedback pitch and time alterations on labeling responses for pitch (**a**) and timing (**b**) of probe events in Experiment 3. *Error bars* reflect one between-participants standard error of the mean
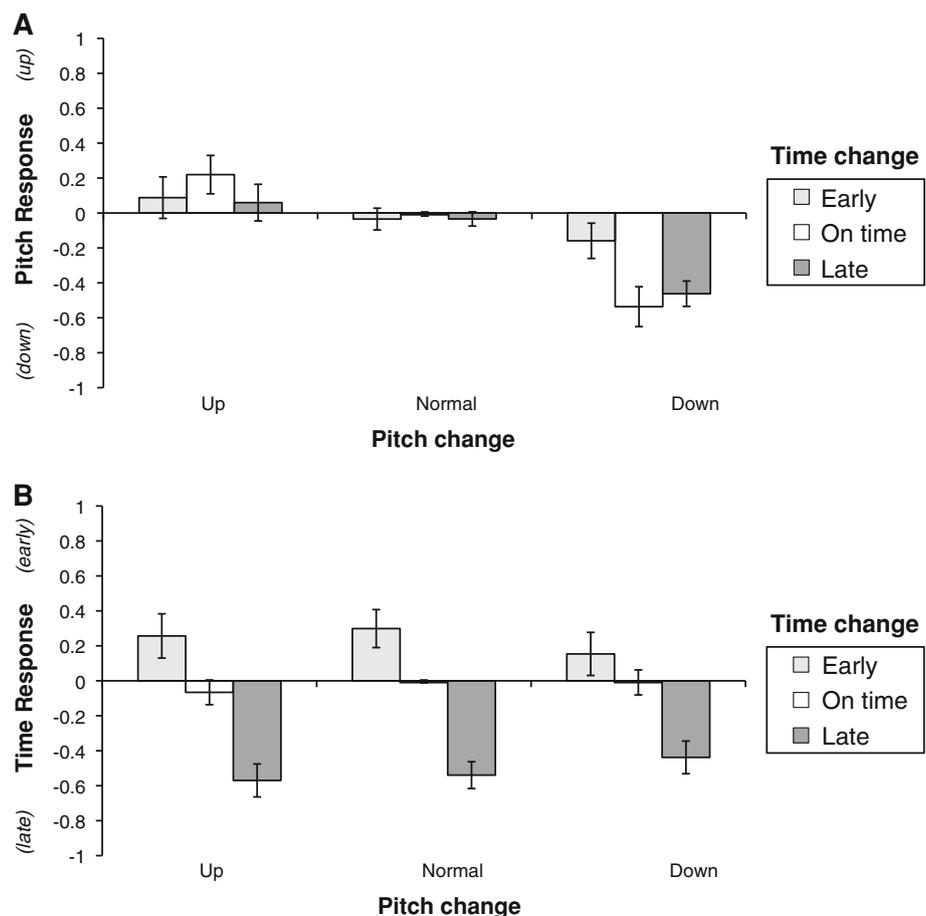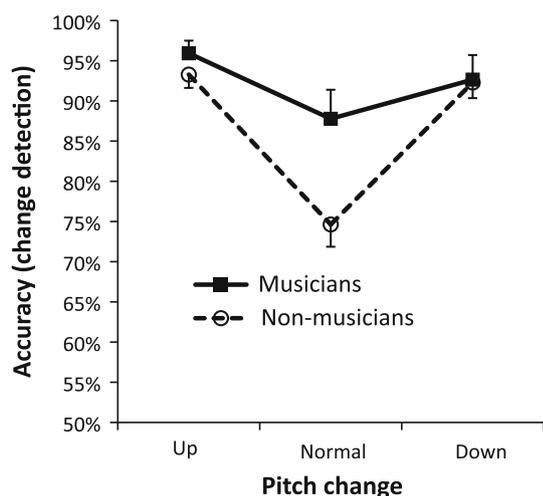
Fig. 8 Significant Experience × Pitch interaction effect for detection of alterations to auditory feedback, pooled across Experiments 1 and 3. *Error bars* reflect one between-participants standard error of the mean

Further analyses explored possible influences of musical training on pitch and time labeling responses. However, no effects of training emerged.

Discussion

Experiment 3 was conducted to test whether the results of Experiment 1 depended on participants' experiencing the standard mapping of responses on the keyboard to the pitch contents of auditory feedback, by reversing this relationship among pitches used to produce melodies. Participants practiced with normal feedback, in order to ensure that they consolidated the standard mapping of responses to effects. Results replicated Experiment 1. This suggests that differences across Experiments 1 and 2 were not simply due to the fact that responses were dissociated from their effects in Experiment 2. Of course, it is possible that participants were able to re-map their conceptualization of responses to effects in Experiment 3 and thus experienced the reversed mapping as if it were normal. However, other, more extreme, forms of perception–action dissociation can lead to disruption of performance, which would have complicated the interpretation of the present results. In this context, it is interesting to recall that a small portion of the participants ($n = 3$) did appear to find the re-mapping quite disruptive (so much so that their data needed to be dropped) and that two of these participants were pianists. Though it is impossible to draw firm conclusions about this unexpected finding, it may reflect greater flexibility among non-pianists when faced with a reversed pitch mapping on the keyboard (cf. Jäncke, 2002).

General Discussion

The three experiments reported here document how well participants can detect alterations of auditory feedback for individual events during performance. Such perturbations have been used in previous studies (Furuya & Soechting, 2010; Maidhof et al., 2010; Pfordresher & Kulpa, 2011). However, this is the first investigation to my knowledge to address the conscious experience of these perturbations during sequence production.

Results suggest that producing a sequence may make one "deaf" to auditory feedback whose pitch contents are associated with the intended pitch at a given serial position, much like planning an action has been shown to produce "blindness" for response-congruent stimuli presented visually in discrete-response tasks (Müsseler & Hommel, 1997a, b; Stoet & Hommel, 1999). These effects may be seen as analogous because the unaltered feedback event maps on to the spatial location of the planned action (a key on the keyboard), similar to how visual stimuli in other research (e.g., an arrow) map on to the spatial coding of a selected action (which hand to use for execution). Importantly, this deafness effect was only found when participants produced motor sequences (via sequential key presses), and was not found when participants generated a melody via isochronous tapping. Thus, it is the generation of a motor plan—associated with the feedback sequence—that leads to response-compatible deafness. Although this effect was specific to sequence production, it was not specific to the particular mapping of keys to pitch on a musical keyboard, given that similar effects were found when the mapping was reversed. Thus, this result shows that earlier results with simpler discrete-response tasks also extend to the production of complex sequences.

At the same time, other effects of timing, for feedback alterations and for production, go beyond results found for discrete-response tasks. First, engaging in a sequence production task selectively disrupts the ability to detect delayed feedback events when pitch is unchanged. This kind of result, reflecting deafness for an event after its production, suggests that temporal binding of actions with events may fluctuate over time and be susceptible to post-output suppression during sequence production. Importantly, post-output suppression is a characteristic that may be necessary for the rapid and fluent production of complex sequences, but may not be necessary for discrete-response tasks. As mentioned in the introduction, other research in musical sequence production has not found evidence for post-output suppression in serial ordering errors (Palmer & Pfordresher, 2003; Pfordresher & Palmer, 2006; Pfordresher et al., 2007). Thus, it is possible that suppression effects only exist on a very fine-grained timescale, as in the asynchronies used

here, and are not detectable when looking at error data. A second effect relates to the influence of overall timing variability on detection of feedback asynchronies: Whereas asynchrony detection was more accurate during isochronous tapping (which was typically produced with higher precision), detection was less accurate during (more temporally variable) sequence production. This result suggests that the salience of asynchronies in performance is influenced by how effectively motor production is entrained to the target tempo. Previous research has shown that a temporally variable auditory precursor sequence degrades the ability to detect asynchronies in a subsequent target sequence (Large & Jones, 1999). The present research suggests such context effects may also be elicited by motor production.

Labeling responses also suggest that engaging in sequence production causes biasing effects of feedback timing on pitch. Specifically, when feedback was presented early, participants had difficulty labeling pitch changes, leading to a null effect of feedback pitch on labeling responses. This result is somewhat similar to the response compatibility blindness effect of Stoet and Hommel (1999) insofar as participants in the present study had trouble comparing presented pitches to the intended pitch content. However, the results of Stoet and Hommel would suggest that deafness effects should be restricted to pitch alterations that go in the direction of the intended pitch. In other words, if the interval preceding the probe tone is an upwards pitch interval, one would expect participants to have a harder time detecting an early pitch that is shifted down relative to the probe pitch than those that are shifted up, given other studies documenting sensitivity of participants in perceptual tasks to pitch/time velocity patterns (e.g., Hubbard, 2005; Johnston & Jones, 2006; Jones, 1976; Jones, Johnston, & Puente, 2006). However, follow-up analyses suggested no difference in pitch labeling response, or accuracy for pitch changes that preserve the trajectory of the sequence versus those that do not (accuracy for trajectory matching changes, $M = 97$ %, SE = 2 %, for trajectory mismatching $M = 97$ %, SE = 2 %; pitch labeling of trajectory matching $M = -.23$, SE = .10, for trajectory mismatching $M = -.18$, SE = .10).

Although the timing of feedback influenced pitch-labeling responses (in Experiments 1 and 3), feedback pitch did not similarly affect the labeling of feedback timing. On the surface, this finding seems to contradict other data suggesting dominance of pitch over time in sequence planning (Prince & Pfordresher, 2012). However, accuracy analyses run counter to this implication in showing better detection of pitch changes than time changes. Ultimately it seems as though the effects of different dimensions on labeling responses reflect the differing roles of pitch and time in this task rather than dominance of one dimension

over the other. Whereas pitch constitutes the content of action planning, timing pertains to the strength of binding between an action plan and auditory feedback (Hommel, 2009). By this logic, the degree of binding may influence one's ability to process content, but the reverse should not occur.

Participants in the experiments reported here varied with respect to musical training, though all performed the same (simple) melodies on a keyboard. Not surprisingly, differing levels of musical experience influenced detection responses. Although no overall effect of training was associated with response accuracy (suggesting that the task was similarly difficult for both groups in general), participants reporting musical training did not exhibit the same feedback deafness for response-compatible delayed feedback, as did non-musicians. This was an unexpected result. However, it is consistent with other data mentioned earlier, suggesting increased sensitivity to altered auditory feedback among pianists, along with analyses of production errors suggesting that piano players do not exhibit post-output suppression. One possible explanation of this effect is that non-musicians approach the present task using the same kind of guiding principles for sequence production that one might use in similar tasks such as speech production, for which post-output suppression is apparent in the error data (Vousden et al., 2000). By contrast, musicians may learn to reduce this suppression given that the frequency of repeated events in music may render post-output suppression disadvantageous (cf. Vousden & Brown, 1998). Obviously, more research is necessary to verify this possibility.

Taken together, the present results suggest that planning motor sequences influences one's ability to process perceptual feedback from those movements. This basic result supports views that perception and action share resources (e.g., Hommel, 2009; Hommel et al., 2001; MacKay, 1987; Shin et al., 2010) and are consistent with computational models proposing that action planning incorporates a "forward model" of perception and action that influences the processing of perceptual feedback (e.g., Jordan & Rumelhart, 1992; Kawato, 1999; Wolpert, 1997). More broadly, the present data accord with other recent findings in suggesting that performers' responses to auditory feedback reflect the interplay of both perception and action planning (cf. Couchman, Beasley, & Pfordresher, 2012).

Appendix

See Tables 1, 2, 3

Table 1 Categorization of labeling response from Experiment 1

	Pitch altered up			Pitch normal			Pitch altered down		
	Early (%)	On time (%)	Late (%)	Early (%)	On time (%)	Late (%)	Early (%)	On time (%)	Late (%)
Pitch response									
Up	**28**	**45**	**30**	18	39	11	26	16	13
Normal	45	11	17	**54**	**21**	**67**	29	8	27
Down	28	45	53	29	40	22	**45**	**76**	**60**
Time response									
Early	**61**	19	16	**58**	19	9	**60**	15	20
On time	9	**60**	17	9	**52**	11	5	**74**	22
Late	30	21	**67**	32	29	**80**	35	11	**58**

Percentages reflect proportion of all trials associated with a "change" response in which the type of change was classified in a certain way. Bold typeface highlights response categories that are correct

Table 2 Categorization of labeling response from Experiment 2

	Pitch altered up			Pitch normal			Pitch altered down		
	Early (%)	On time (%)	Late (%)	Early (%)	On time (%)	Late (%)	Early (%)	On time (%)	Late (%)
Pitch response									
Up	**49**	**58**	**48**	35	32	16	34	21	23
Normal	25	2	16	**43**	**20**	**57**	19	11	17
Down	26	40	36	22	48	26	**47**	**68**	**59**
Time response									
Early	**55**	14	16	**70**	28	11	**66**	14	19
On time	5	**77**	10	3	**54**	5	5	**73**	16
Late	40	9	**75**	27	18	**84**	29	13	**66**

Percentages reflect proportion of all trials associated with a "change" response in which the type of change was classified in a certain way. Bold typeface highlights response categories that are correct

Table 3 Categorization of labeling response from Experiment 3

	Pitch altered up			Pitch normal			Pitch altered down		
	Early (%)	On time (%)	Late (%)	Early (%)	On time (%)	Late (%)	Early (%)	On time (%)	Late (%)
Pitch response									
Up	**43**	**57**	**44**	23	35	17	27	19	16
Normal	28	5	17	**62**	**27**	**68**	33	8	30
Down	29	38	39	15	39	15	**40**	**73**	**54**
Time response									
Early	**54**	10	13	**61**	26	7	**57**	13	16
On time	14	**74**	26	7	**45**	8	10	**71**	24
Late	32	16	**61**	32	29	**85**	33	15	**60**

Percentages reflect proportion of all trials associated with a "change" response in which the type of change was classified in a certain way. Bold typeface highlights response categories that are correct

References

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

Couchman, J. J., Beasley, R., & Pfordresher, P. Q. (2012). Auditory feedback, self-attribution, and the experience of agency in sequence production. *Consciousness and Cognition, 21*, 186–203.

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.

Ellis, R. J., & Jones, M. R. (2009). The Role of Accent Salience and Joint Accent Structure in Meter Perception. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 264–280.

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

Furuya, S., & Soechting, J. (2010). Role of auditory feedback in the control of successive keystrokes during piano playing. *Experimental Brain Research, 204*, 223–237.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review, 102*, 594–621.

Hébert, S., & Peretz, I. (1997). Recognition of music in long-term memory: Are melodic and temporal patterns equal partners? *Memory & cognition, 25*, 518–533.

Hommel, B. (2009). Action control according to TEC (theory of event coding). *Psychological Research, 73*, 512–526.

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

Howell, P. (2001). A model of timing interference to speech control in normal and altered listening conditions applied to the treatment of stuttering. In B. Maassen, W. Hulsijn, R. Kent, H. F. M. Peters & P. H. M. M. van-Lieshout (Eds.), *Speech motor control in normal and disordered speech* (pp. 291–294). Nijmegen: Uttgeverij Vantilt.

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 and Performance, 9*, 772–784.

Hubbard, T. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review, 12*, 822–851.

Jäncke, L. (2002). The case of a left-handed pianists playing a reversed keyboard: A challenge for the neuroscience of music. *NeuroReport, 13*, 1579–1583.

Johnston, H. M., & Jones, M. R. (2006). Higher order pattern structure influences auditory representational momentum. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 2–17.

Jones, M. R. (1976). Time, our lost dimension—Toward a new theory of perception, attention, and memory. *Psychological Review, 83*, 323–355.

Jones, M. R., Johnston, H. M., & Puente, J. (2006). Effects of auditory pattern structure on anticipatory and reactive attending. *Cognitive Psychology, 53*, 59–96.

Jordan, M. I., & Rumelhart, D. E. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive Science, 16*, 307–354.

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Biology, 9*, 718–727.

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

Keppel, G., & Wickens, T. D. (2004). *Design and analysis: A researcher's handbook* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin, 126*, 159–179.

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

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*, 119–159.

Lashley, K. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral Mechanisms in Behavior* (pp. 112–136). New York: Wiley.

Lidji, P., Kolinsky, R., Lochy, A., & Morais, J. (2007). Spatial associations for musical stimuli: A piano in the head? *Journal of Experimental Psychology: Human Perception and Performance, 33*, 1189–1207.

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

Maidhof, C., Vavatzanidis, N., Prinz, W., Rieger, M., & Koelsch, S. (2010). Processing Expectancy Violations during Music Performance and Perception: An ERP Study. *Journal of Cognitive Neuroscience, 22*, 2401–2413.

Müsseler, J., & Hommel, B. (1997a). Blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 861–872.

Müsseler, J., & Hommel, B. (1997b). Detecting and identifying response-compatible stimuli. *Psychonomic Bulletin & Review, 4*, 125–129.

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, 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.

Pfordresher, P. Q. (2003). Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing. *Journal of Experimental Psychology: Human Perception and 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 and Performance, 31*, 1331–1345.

Pfordresher, P. Q. (2006). Coordination of perception and action in music performance. *Advances in Cognitive Psychology. Special Issue: Music performance, 2*, 183–198.

Pfordresher, P. Q. (2012). Musical training and the role of auditory feedback during performance. *Annals of the New York Academy of Sciences, 1252*, 171–178.

Pfordresher, P. Q., & Dalla-Bella, S. (2011). Delayed auditory feedback and movement. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 566–579.

Pfordresher, P. Q., & Kulpa, J. D. (2011). The dynamics of disruption from altered auditory feedback. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 949–967.

Pfordresher, P. Q., & Palmer, C. (2006). Effects of hearing the past, present, or future during music performance. *Perception and Psychophysics, 68*, 362–376.

Pfordresher, P. Q., Palmer, C., & Jungers, M. K. (2007). Speed, Accuracy, and Serial Order in Sequence Production. *Cognitive Science: A Multidisciplinary Journal, 31*, 1–36.

Prince, J. B., & Pfordresher, P. Q. (2012). The role of pitch and temporal diversity in the perception and production of musical sequence. *Acta Psychologica, 141*, 184–198.

Prince, J. B., Schmuckler, M. A., & Thompson, W. F. (2009a). The effect of task and pitch structure on pitch-time interactions in music. *Memory and Cognition, 37*, 368–381.

Prince, J. B., Thompson, W. F., & Schmuckler, M. A. (2009b). Pitch and time, tonality and meter: How do musical dimensions combine? *Journal of Experimental Psychology: Human Perception and Performance, 35*, 1598–1617.

Rusconi, E., Kwan, B., Giordano, B. L., Umilta, C., & Butterworth, B. (2006). Spatial representation of pitch height: the SMARC effect. *Cognition, 99*, 113–129.

Shin, Y. K., Proctor, R. W., & Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychological Bulletin, 136*, 943–974.

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

Stoet, G., & Hommel, B. (1999). Action planning and the temporal binding of response codes. *Journal of Experimental Psychology: Human Perception and Performance, 25*.

Vousden, J. I., & Brown, G. D. A. (1998). To repeat or not to repeat: The time course of response suppression in sequential behaviour. In J. A. Bullinaria, D. W. Glasspool, & G. Houghton (Eds.), *Proceedings of the Fourth Neural Computation and Psychology Workshop: Connectionist Representations*. London: Springer.

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.

Wheeldon, L. R., Meyer, A. S., & Smith, M. (2002). Incrementality. In R. Goldstone (Ed.), *Encyclopedia of cognitive science*. London: Macmillan.

Wolpert, D. M. (1997). Computational approaches to motor control. *Trends in Cognitive Sciences, 1*, 209–216.