RESEARCH ARTICLE

Effects of delayed auditory and visual feedback on sequence production

J. D. Kulpa · Peter Q. Pfordresher

Abstract This study represents the first systematic comparison of the relative contributions of auditory and visual feedback to sequence production. Participants learned an isochronous melody that they performed on a keyboard and attempted to perform this sequence at a prescribed rate while auditory and visual feedback were manipulated. Delayed auditory feedback (DAF) and delayed visual feedback (DVF) both tended to slow production of the sequence. These effects were additive. There was no modulation of this effect of delay in either modality by the absence of feedback in the other. In contrast with past research, DAF did not increase timing variability, though DVF did. Motion analyses ruled out differences in salience of visual feedback between delayed and non-delayed conditions as an explanation of the effects of DVF. The results suggest that the effects of delayed feedback may be attributable to both sensorimotor interference and to conflicting information across feedback channels.

Keywords Sequence production · Timing · Auditory feedback · Visual feedback · Delayed feedback

J. D. Kulpa
Department of Psychology, New Mexico State University,
317 Science Hall, Las Cruces, NM 88003, USA

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

Introduction

The actions people produce are typically accompanied by correlated perceptual events, referred to as perceptual feedback. It has been suggested that performers plan actions by generating a mental image of the anticipated perceptual consequences of that action (e.g., Hommel et al. 2001; James 1890; Prinz et al. 2009). If so, action planning should be sensitive to mismatches between the anticipated and actual outcomes of actions (Keller and Koch 2008). Indeed, alterations of perceptual feedback can disrupt the execution (and possibly also the planning) of actions. It is well known that delayed auditory feedback (DAF) during music performance (and speech) can disrupt production, primarily by causing timing to slow down and become more variable (for reviews see Howell 2001; Pfordresher 2006; Yates 1963). Delayed visual feedback (DVF) has been found to diminish accuracy during tracking tasks (e.g., Langenberg et al. 1998; Smith et al. 1960) and to destabilize postural control (van den Heuvel et al. 2009). However, to our knowledge, nobody has systematically investigated the separate and joint effects of DAF and DVF on sequence production. We here consider a performance domain—playing a melody on a keyboard—in which the performer intends to reproduce a sequence of learned finger movements at a given rate, while auditory feedback is presented over headphones and visual feedback is presented as a motion capture animation of the performer's hand. For the sake of simplicity in this initial investigation, all manipulations of feedback timing, both auditory and visual, are set at a 300 ms delay. As we summarize below, systematically varying the auditory and visual feedback conditions under which participants carry out this task allows us to test several hypotheses concerning the effect of delayed feedback.

The sensorimotor conflict hypothesis

We first consider the most common interpretation of disruption from delayed feedback, the sensorimotor conflict hypothesis. It says that delayed feedback interferes with the planned timing of actions (Pfordresher 2006) or their execution (Howell 2001) because of shared representations for perception and action (e.g., Hommel et al. 2001; MacKay 1987). For example, the Theory of Event Coding (Hommel et al. 2001) suggests that action plans are coded in terms of the anticipated distal consequence of those actions. Delayed feedback is thus disruptive because its timing conflicts with planned movement timing (Pfordresher and Dalla Bella 2011).

The sensorimotor conflict hypothesis predicts that the presence of either DAF or DVF alone could disrupt performance. However, the degree of disruption by a single delayed source of feedback is likely to depend on the perceptual salience of the delayed feedback. The simultaneous presence of DAF and DVF should be more disruptive than the presence of just one of them, and we might expect their effects to be additive.

The feedback conflict hypothesis

Another possible interpretation of disruption from delayed feedback focuses on interference between feedback of different modalities, instead of interference between anticipated and actual feedback events. Perceptual feedback is transduced along multiple sensory channels, including auditory, visual, tactile, and proprioceptive modalities. When feedback within one modality is delayed (as in DAF, for example), the result is conflicting temporal information across feedback channels. Thus, the disruptive effect of manipulations like DAF may, at least in part, be based on conflicts between perceptual channels, rather than resulting exclusively from conflicts between anticipated and perceived feedback.

The current experiment was designed to test the prediction that disruption from altered feedback scales with the degree of conflict across perceptual channels. Two extremes may be considered: When all feedback channels are synchronized (either because feedback has not been altered or because feedback has been delayed to the same degree), there is no conflict, and thus, there should be no disruption according to the hypothesis. By contrast, maximal conflict (and maximal disruption) should occur when approximately half of the available channels are delayed and the rest are synchronized with action. This prediction is made under the simplifying assumption that the potential to disrupt sequence production is the same for conflict in any given pair of feedback channels. Under this assumption, degree of disruption can be estimated by comparing the number of pairs of mismatched feedback sources with the number of pairs of matched sources. Assume we start with multiple feedback sources synchronized with action (e.g., proprioceptive, tactile, and auditory) and one delayed feedback source (e.g., visual). In this scenario there are three mismatched pairs and three matched pairs. If we then remove one of the non-delayed sources (e.g., auditory), the disruptive effect of the DVF should increase because, with two mismatched pairs and only one matched pair, that delayed source now constitutes a larger proportion of all available channels.

The feedback conflict hypothesis is not easy to test, however. To test this hypothesis with regard to auditory and visual feedback, we would ideally have to remove all other sources of feedback and then compare the disruptive effects of having one and both of these sources delayed to a baseline of normal feedback only. Unfortunately, it is impossible (as far as we know) to remove all feedback other than auditory and visual (even if we were to remove tactile feedback, for instance, participants would still have access to proprioception). In the current experiment, we tested the feedback conflict hypothesis by presenting DAF and DVF in trials where feedback from the other modality may be normal, absent, or delayed by an equal amount. First, the hypothesis predicts that, when no modalities are delayed, performance should not be affected when feedback in one or two modalities is removed. However, the effect of delayed feedback from either the auditory or visual modality should increase when the other modality is absent relative to when the other modality is normal and should be greatest when both auditory and visual modalities are delayed. Notice, though, that the latter would not necessarily be the case if auditory and visual feedback were delayed by unequal amounts or if the assumption of equal disruptive potential across feedback channel pairs did not hold. For instance, if auditory and visual feedback were especially important in comparison with tactile and proprioceptive feedback, then having both auditory and visual feedback delayed by the same amount would result in less disruption than if only one of those delayed sources were present.

The auditory dominance hypothesis

Several studies indicate that the auditory modality dominates the visual modality with respect to the encoding of purely temporal information used to time actions. Participants demonstrate lower precision when synchronizing taps with visually presented rhythms than auditory rhythms (Kolers and Brewster 1985; Repp and Penel 2002, 2004) and are less sensitive to perturbations in the visual than the auditory rhythm (Repp and Penel 2002, 2004). When auditory and visual stimuli are presented simultaneously,

participants respond almost exclusively to temporal information in the auditory modality (Repp and Penel 2002, 2004). One limitation of these past studies is that in each case, visual information was presented discretely, as flashes of light. By contrast, temporal information in the visual modality is more commonly associated with continuous motion (cf. Gibson 1966; Lee and Reddish 1981). Thus, it is possible that the dominance of the auditory modality found in previous research may reflect the limitations of the visual system in processing temporal information related to abrupt, discrete events (cf. Hove et al. in press). In this context, the use of DVF is convenient, in that delays of visual feedback from continuous movements may provide more salient temporal information than the discrete visual events in previous studies concerning synchronization.

Thus, we used the present paradigm as a way of revisiting the auditory dominance hypothesis in a different context—one in which visual feedback is continuous. Based on the results of the most closely related previous study, by Repp and Penel (2004), we hypothesized that auditory dominance should lead to negligible effects of visual feedback, except perhaps when auditory feedback is absent. In other words, the disruptive effect of delayed feedback may be specific to the auditory modality, possibly relating to the tight coupling between auditory and tactile feedback systems (Occelli et al. 2011). Of course, it is also possible that auditory and visual information may simply differ with respect to the magnitude of their effects (a conservative version of the auditory dominance hypothesis).

In the present study, we did not attempt to calibrate the salience of auditory and visual delays for participants. Though this is a limitation, previous research has demonstrated persistence of auditory dominance even when the intensity of the auditory stimuli is substantially diminished while the intensity of the visual stimuli is held constant, possibly because salience of auditory temporal information does not depend on intensity, as long as intensity is above threshold (Repp and Penel 2004). Because it is difficult to perfectly match salience of auditory and visual feedback, we test a strong version of the auditory dominance hypothesis, namely that DVF will have no reliable effect on performance when combined with auditory feedback, but will have a reliable effect when auditory feedback is absent.

The present experiment

In the present experiment, participants experienced auditory and visual feedback that was either normal, absent, or delayed by 300 ms, in a factorial design. A critical question concerns the effects of removing one feedback modality (auditory or visual) on delays within the other modality. Whereas the sensorimotor conflict hypothesis predicts that removing feedback from one modality will have no effect on delays in the other modality, the feedback conflict hypothesis predicts that the disruptive effect of delays from one modality should increase when feedback from the other modality is removed. Furthermore, the auditory dominance hypothesis predicts that visual delays will have little effect on performance when auditory feedback is present, though visual delays presented with auditory feedback absent might still disrupt performance (albeit to a lesser degree than auditory delays).

Method

Participants

Twelve individuals (mean age = 21, range = 18–26) volunteered to participate in exchange for Introductory Psychology course credit at the University at Buffalo, State University of New York. Three participants were female and nine were male. All participants reported being right-handed. Five participants reported one or more years of musical training ($M = 7.8$ years, range = 1–23), but only one reported piano training (10 years, this was the same participant who reported 23 total years of musical training). Results from the highly trained participant resembled those of other participants; thus, this participant was not a statistical outlier with respect to the effects of feedback manipulations. No participants reported having absolute pitch. Participants reported no hearing impairment and no motor impairment of the right hand.

Apparatus

Participants played on an M-AUDIO Keystation 49e unweighted piano keyboard. Auditory feedback manipulations were made, and the timing of keystrokes and feedback recorded, using the program FTAP (Finney 2001). Participants heard auditory feedback and metronome pulses over Sony MDR-7500 professional headphones at a comfortable listening level. Auditory output was produced by a Roland RD-700 digital piano. The timbre of keystrokes was set to Program 1 (Standard Concert Piano 1) of the RD-700, and the metronome timbre was set to Program 126 (standard set, MIDI key 56 = cowbell).

Visual feedback was produced using a Visualeyez single-tracker active motion capture system (Phoenix Technologies, Burnaby, BC, Canada). Participants had 11 markers placed on their right hand (one on each fingernail, one on each base knuckle, and one at the wrist). Four markers were also attached to the corners of the keyboard. Through the use of customized software (VZAnalyzer), a continuously updated stick-figure image of their hand and
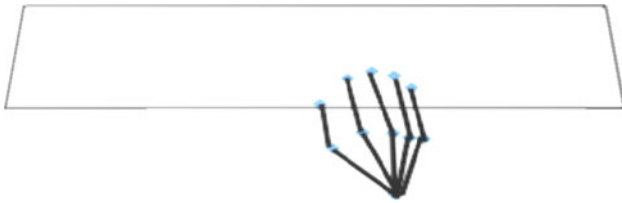
Fig. 1 Reconstruction of the hand image displayed to participants as visual feedback. See the text for details

an image of the keyboard as a solid colored plane were presented to participants on a computer monitor. Figure 1 shows a representation of the hand and keyboard similar to what participants would see while playing. A music stand was positioned to act as an occluder, such that, when seated in a comfortable position to play, participants could easily see the computer image directly in front of them but could not see their hand. Delayed visual feedback of the image was presented using customized software (VZSoft), and the absence of visual feedback was accomplished by minimizing the VZAnalyzer window, revealing a black screen. The temporal resolution of the streamed video was tested by analyzing a video recording of hand movements combined with synchronized visual feedback frame-by-frame. These analyses suggested that the streamed video had a transmission delay of 6 ms, a shorter delay than can often be found for the timing of auditory events via MIDI (approximately 20 ms for the apparatus used here). Time lags of this magnitude are not known to cause disruption, though they may have subtle effects on the timing of production (Madison and Merker 2004).

Conditions

The experiment had a two-factor repeated-measures design. Nine experimental conditions were produced by fully crossing the factors auditory feedback (normal, absent, delayed) with visual feedback (normal, absent, delayed). When auditory feedback was absent, no sound was presented over headphones; when visual feedback was absent, the screen was blank. When auditory feedback was delayed, MIDI tone onsets lagged after their associated keypress by 300 ms (a delay amount likely to cause disruption), and when visual feedback was delayed, the streamed animation of the participant's hand was delayed throughout the trial by 300 ms.

Each condition was repeated 3 times, leading to 27 trials in a session. Trials were divided into 3 blocks, each of which comprised a different random order of the 9 trial types. Participants were randomly assigned to one of two random orders of all trials (within the constraints described above).

Materials

All participants performed the same 8-note musical sequence, C4-D4-E4-G4-F4-E4-D4-E4 (where the letter represents the pitch class and the number represents the octave), used in previous research (Pfordresher 2005, 2008). Music was displayed using notation that is designed for participants who are not musically trained (first described in Pfordresher 2005). Notes were displayed as a numerical sequence [1 2 3 5 4 3 2 3] below a series of right-hand images with the appropriate finger highlighted (in this case: thumb = 1, index finger = 2, etc.). A fixed finger-key mapping was used so that participants never had to shift their hand. Above the keyboard were labels showing the numbers 1–5 with arrows pointing to the corresponding keys.

Procedure

After the requirements of the experiment were explained and informed consent was obtained, an outline of participants' right hand was traced with the location of the markers indicated. Recording the placement of the motion capture markers on this tracing helped to ensure placement was consistent across participants. Eleven motion capture markers were then affixed to the hand in the locations described above.

Participants maintained a comfortable seated position at the keyboard while the occluder was set up. Musical notation was affixed to the computer screen directly in front of participants. During an initial memorization phase, participants received no visual feedback and normal auditory feedback. First, the experimenter had participants play through the 8-note sequence once, in order to be sure that they understood the notation. Participants then practiced playing the melody until they thought that it was learned. The experimenter then removed the musical notation sheet, and participants continued to play. If it was not clear that the sequence had been memorized, the music was returned to view and practice continued. When participants played the melody three times through error-free without the music, the experiment moved into the practice trial phase.

Participants were instructed to always watch the screen when a trial was underway, though they were not explicitly instructed to attend to any feedback source. The experimenter monitored participants to ensure that they were oriented toward the screen during trials. When visual feedback was normal or delayed, participants were instructed to watch the displayed image of their hand. When visual feedback was absent, participants were instructed to look at the blank black screen. Each trial began with four metronome pulses (abrupt percussive sound resembling a cowbell) presented with an inter-onset

interval of 500 ms. The metronome was only present for these first four beats. Participants were instructed to listen to these pulses, and then to begin playing the melody on the fifth beat, attempting to replicate the pace of the metronome. Participants were instructed to continue performing the melody without stopping, even when disruption occurred. Each trial began with the four metronome pulses and ended when the experimenter said, "stop." Participants were instructed to simply play the melody repeatedly until they heard this verbal cue. Every trial (practice and experimental) consisted of 48 keystrokes (six repetitions of the melody).

Prior to the experimental trials, participants completed four practice trials designed to familiarize them with manipulations of auditory and visual feedback. These four practice trials were (in order) (1) normal auditory/normal visual, (2) delayed auditory/normal visual, (3) normal auditory/delayed visual, and (4) delayed auditory/delayed visual. Participants were not exposed to the feedback absent conditions during practice but were informed that for some trials the screen would be blank and/or nothing would be heard over the headphones. After the practice trials were completed, experimental trials (as described above) began. When participants had completed all experimental trials, they were asked to complete a series of questionnaires concerning their musical experience, hearing acuity, and demographic information. Finally, participants were debriefed and thanked.

Data analysis

We extracted inter-response intervals (IRI = time from one keypress to the next, in ms) from each trial to measure disruption of timing after removing timing outliers (defined as IRIs outside a range of ±2 standard deviations for that trial). In addition, we removed variance associated with tempo drift for each trial by taking the residuals of a linear regression of IRI on serial position and adding the mean IRI for the trial to those residuals (as a result, variability associated with the linear trend is removed but the mean for the trial remains the same). We analyzed two measures of timing disruption: the mean of IRIs in a trial, which is typically larger (indicating a slower tempo) during disruptive conditions, and the coefficient of variation (CV = SD/mean) of IRIs, which is also higher when participants are disrupted while attempting to perform isochronously. Participants made almost no errors during the experiment (which is not surprising for the altered auditory feedback conditions used here, cf. Pfordresher 2003; Pfordresher and Kulpa 2011), so we did not analyze error rates. Each measure of timing was analyzed using a 3 (auditory feedback) × 3 (visual feedback) repeated-measures analysis of variance (ANOVA).
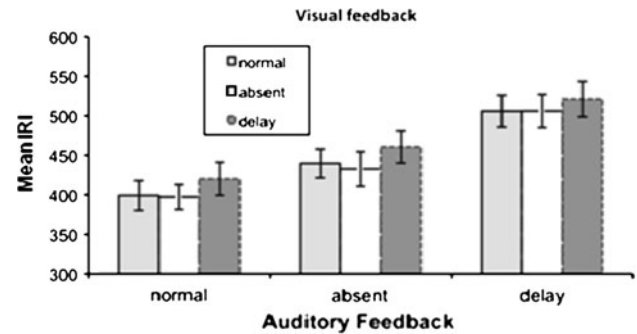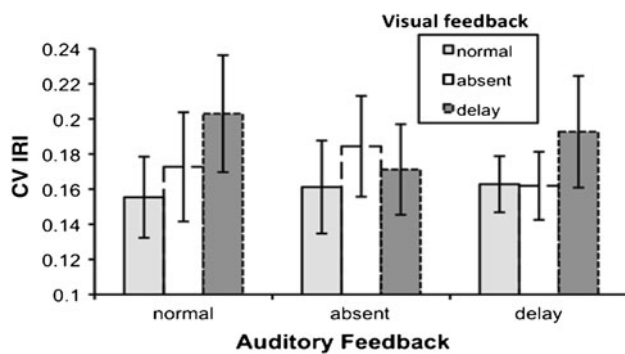

Fig. 2 Mean inter-response intervals (IRIs) across participants for each of the nine feedback conditions. *Error bars* display one standard error of the mean

Results

Figure 2 shows mean IRIs across all nine feedback conditions. During trials with normal feedback, participants performed considerably faster than the target rate of the metronome. In similar research, participants typically perform faster than the metronome rate after the metronome stops (cf. Pfordresher 2003, 2005), a tendency that was probably exaggerated here because participants did not explicitly synchronize with the metronome. The ANOVA yielded a main effect of auditory feedback, $F(2, 22) = 61.6$, $p < .01$, $\eta_p^2 = .84$, a main effect of visual feedback, $F(2, 22) = 8.24$, $p < .01$, $\eta_p^2 = .43$, but no interaction ($F < 1$, $p = .70$, $\eta_p^2 = .05$). Post hoc contrasts were performed on the main effects of auditory and visual feedback, using the Sidak–Bonferroni correction for contrasts within each effect (pairwise corrected $\alpha = .02$). For the main effect of auditory feedback, DAF led to significantly higher IRIs ($M = 514$, SE $= 11.9$) than normal auditory feedback ($M = 407$, SE $= 10.8$), $t(11) = 10.50$, $p < .01$, and absent feedback ($M = 444$, SE $= 11.2$), $t(11) = 7.87$, $p < .01$. In addition, IRIs for absent feedback were significantly longer than IRIs for normal feedback, $t(11) = 4.24$, $p < .01$. For the main effect of visual feedback, IRIs were longer for DVF ($M = 467$, SE $= 13.1$) than for trials with normal visual feedback ($M = 448$, SE $= 13.1$), $t(11) = 3.57$, $p < .01$, and absent visual feedback ($M = 450$, SE $= 13.7$), $t(11) = 3.57$, $p < .01$. Conditions with absent versus normal visual feedback did not differ from each other, however.

We also analyzed the effect of DAF and DVF on timing variability in production; mean CVs of IRIs are shown in Fig. 3. In contrast to the results for mean IRI, the only significant effect was the main effect of visual feedback, $F(2, 22) = 4.37$, $p = .03$, $\eta_p^2 = .28$, the main effect of auditory feedback was not significant ($F < 1$, $p = .93$, $\eta_p^2 = .01$), nor was the interaction ($F < 1$, $p = .54$, $\eta_p^2 = .07$). Post hoc contrasts on the main effect of visual

Fig. 3 Mean coefficients of variation (CV IRI = SD/M) across participants for each of the nine feedback conditions. *Error bars* display one standard error of the mean

feedback (conducted as for mean IRI data) only yielded a significant difference between DVF and normal visual feedback, $t(11) = 2.93$, $p < .01$.

Next we report two analyses that were performed to determine whether variations in the salience of visual rhythm caused by finger movements may have influenced performance in selected conditions. Specifically, it is plausible that differences in the precision of finger movements, and the range of motion used, during delayed conditions may have influenced the salience of visual versus auditory delays. One analysis we used to address precision in movements (as opposed to the precision of key press times) was to submit movement position data (the Z coordinate of motion data) to a Fourier analysis, using the fast Fourier transform function of Matlab (Mathworks, Natick, MA, USA). The resulting power spectrum of this transform yields peaks that represent dominant periodicities in finger movements, whereas the degree of spread around these peaks represents the reliability (precision) of these periods, which is our primary concern here. Figure 4 shows the power spectrum from this analysis for conditions with normal or delayed auditory feedback (Fig. 4a), averaged across all visual feedback conditions, and normal versus delayed visual feedback (Fig. 4b), averaged across all auditory feedback conditions. Though power spectra were computed separately for each finger, we present averages across fingers because we are interested in the effects of feedback condition rather than finger-specific effects. The dominant frequency in conditions with normal feedback (apparent in both panels) was .33 Hz, a period of slightly over 3 s. Given the mean IRI for normal feedback (399 ms), this period reflects finger movements approximately every 8 key presses, which was the period of finger movements for 3 out of 5 fingers, the ones that were used only once during each playing of the 8-note melody (thumb, ring, and pinky, see "Materials"). The fact that this peak is shifted to the left for DAF trials (Fig. 4a) simply reflects the slowing effect shown in Fig. 2.

Our primary concern here was whether the spread around this peak varies with condition. In a power spectrum, the degree of spread around a frequency reflects how dominant that frequency is; in the present context, it reflects the precision of movements with respect to dominant periods (by analogy, in signal processing a pure tone has a maximally precise representation of frequency). The precision of visual rhythmic motion (visually apparent from the distribution in each data series around the modal frequency) was similar across conditions, and therefore, effects of delayed feedback are unlikely to derive from differences in temporal movement precision. The same conclusions held for power spectra within each finger, even though differences in biomechanical constraints and frequency of use in the sequence led to finger-specific differences in spectral shape.

Our second movement-related analysis focused on changes in finger height across trials.[1] Finger height can increase with performance tempo (Dalla Bella and Palmer 2011), can be used to enhance tactile feedback (Goebl and Palmer 2008), and—most important for the present study—can influence the perceptual salience of visual motion. We analyzed finger movements by extracting the maximum and minimum values for each finger in a trial, leading to a difference score that reflects the maximum range of motion, which was then averaged across trials within a condition for each participant. These data were analyzed using a 3-way ANOVA with the factors auditory feedback, visual feedback, and finger. Mean range of finger motion as a function of auditory and visual feedback is shown in Fig. 5. The ANOVA yielded a significant interaction between auditory and visual feedback conditions, $F(4, 44) = 2.97$, $p = .03$, $\eta_p^2 = .21$, but no main effect of auditory feedback ($F < 1$, $p = .57$, $\eta_p^2 = .05$) or visual feedback ($F < 1$, $p = .57$, $\eta_p^2 = .05$). There was a significant main effect of finger, $F(4, 44) = 6.75$, $p < .01$, $\eta_p^2 = .38$, with larger ranges of motion for the index and middle finger than for other fingers, but this did not interact with either feedback condition. Moreover, the interaction shown in Fig. 5 does not follow a pattern that resembles the influence of feedback condition on either timing measure. Thus, although finger height varied across conditions it did not do so in a way that confounded the effect of auditory or visual feedback on the timing of onsets.

Discussion

The current experiment is the first to our knowledge to test the relative contributions of auditory and visual feedback to

---

[1] Differences in mean finger length across participants were very small (SD = .43 cm) and were not significantly correlated with the effects of DAF or DVF.

Fig. 4 Power spectra associated with normal versus DAF conditions (**a**), and normal versus DVF conditions (**b**). Spectra shown here are averaged across fingers, trials, participants, and all levels of the feedback factor not shown (e.g., Fig. 4a is averaged across all visual feedback conditions). *Vertical bars* highlight divisions of the beat (period = 500 ms). See text for further details
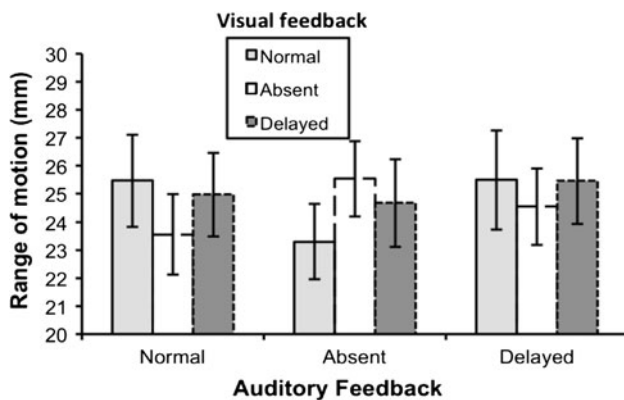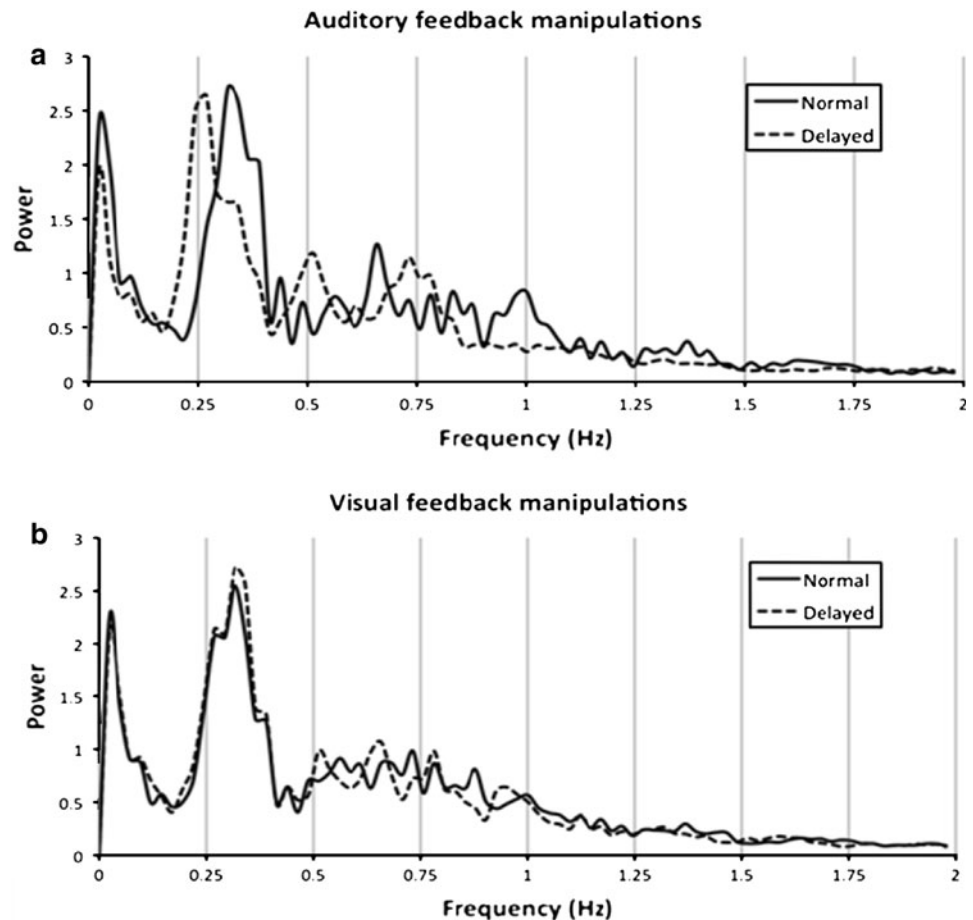
Fig. 5 Mean peak range of motion averaged across fingers trials and participants, for each of the nine feedback conditions. *Error bars* display one standard error of the mean

the timing of musical sequence production. We found that DAF and DVF independently contributed to slowing the production of learned sequences of finger movements. In addition, absent auditory feedback slowed timing somewhat, though the absence of visual feedback had no effect on production rate. Interestingly, we found only an effect of visual feedback on timing variability, with timing variability increasing for trials with DVF versus normal visual feedback. This contrasts with other studies (varying in degree of expertise possessed by participants) that have shown significant influence of DAF on timing variability (e.g., Howell and Sackin 2002; Pfordresher and Palmer 2002; Pfordresher and Dalla Bella 2011). The effects we observed were not a by-product of the precision or amplitude of visually presented hand movements. Thus, we interpret these results to reflect disrupted synchronization of the timing of movements with the periodic rhythmic information presented via the auditory or visual modalities. One plausible explanation for the slowing effect of delayed feedback is that participants are attempting to synchronize their actions with that delayed feedback, as in pseudo-synchronization (Flach 2005).

We designed this experiment to test three hypotheses regarding the disruptive effect of delayed feedback. Past theoretical accounts have tended toward explanations based on motor planning and execution, or sensorimotor conflict (Howell 2001; MacKay 1987; Pfordresher 2006). By contrast, the feedback conflict hypothesis suggests that disruption is the result of timing mismatches between sensory modalities, rather than between actual and

anticipated perceptual outcomes. Finally, the auditory dominance hypothesis stemmed from past research suggesting that, with stationary stimuli, the auditory modality conveys temporal information more effectively than the visual modality. None of these three hypotheses was fully supported. The partial support each received is discussed below. These mixed results suggest that the effect of altered feedback may be based on a combination of motor and perceptual factors. This conclusion is similar in spirit to an application of the common coding approach to the effect of DAF, which considers how participants compare auditory and tactile inputs while synchronizing with a metronome (Aschersleben and Prinz 1997). However, we know of no model (including common coding) that predicts the specific pattern of results reported here.

A critical prediction of the feedback conflict hypothesis for the present study was that the effect of delayed feedback in one modality should increase when information in another modality is removed. However, the effect of auditory feedback absence was to slow down performances across all visual feedback conditions, including performances with normal or absent visual feedback. Removal of visual feedback had no apparent effects on mean IRI. For timing variability, there was an apparent reduction in the effect of DVF when auditory feedback was absent, but this influence was unreliable. The second prediction of the feedback conflict hypothesis, that delays in two modalities would lead to a stronger effect than a delay in one modality, was supported. This finding may be taken as partial support for the feedback conflict hypothesis, in that it suggests that the effect of delayed feedback increases when the ratio of delayed to normal feedback channels approaches 1:1. However, this result does not refute the traditional interpretation of disruption from altered feedback as resulting from interference between perception and action (i.e., the sensorimotor hypothesis), since both hypotheses make the same prediction in this circumstance.

The auditory dominance hypothesis suggests an interaction, such that visual feedback influences production when auditory feedback is absent, but not when auditory feedback is present. Contrary to this prediction, no interaction was found and DVF slowed production relative to normal visual feedback across all levels of auditory feedback. Of course, the fact that visual feedback yielded a smaller effect on mean IRI than did auditory feedback appears to support a more conservative version of the auditory dominance hypothesis (i.e., one that simply predicts a greater effect of auditory than visual feedback). However, we also found that visual rather than auditory feedback influenced timing variability, thus arguing for visual dominance with respect to the precision of produced timing. Taken together, the strong version of the auditory dominance hypothesis is not supported by the present data.

In a practical sense, the relative weakness of visual feedback here may reflect the fact that piano performers (unlike in the present experiment) often do not look at their fingers as they perform.

We suspect that the current results differ from past results that have shown dominance of the auditory modality in sensorimotor tasks because visual feedback in the current study was continuous rather than discrete (as in Kolers and Brewster 1985; Patel et al. 2005; Repp and Penel 2002, 2004). An interesting but unexpected result emerged in the apparently greater effect of visual than auditory information on timing variability. Other recent data suggest that perceptual feedback influences the rate and precision of timing in different ways (Pfordresher and Dalla Bella 2011), and it is possible that these aspects of timing are differentially sensitive to information in the auditory and visual modalities. Production rate and precision may also be differentially sensitive to continuous versus discrete presentations of perceptual feedback. Specifically, alterations of discretely presented feedback may tend to affect rate, whereas precision may be affected more by changes in continuous feedback. In this experiment, the visual presentation of feedback provided enough detail (and the delay was long enough) that, in the delayed condition, participants often perceived the movement of a finger different from the one they were actually moving at the time. This is a possible explanation for the visual feedback effect on precision. Clearly, more experimentation is necessary to conclusively determine whether differential sensitivities of rate and timing to auditory versus visual or discrete versus continuous feedback exist.

A limitation of the current paradigm is the use of a simplified (stick) image of the hand for visual feedback. Another limitation is that, while auditory feedback was provided in a fairly typical manner (i.e., sound over headphones, while playing on a keyboard), visual feedback was provided in a quite unusual manner (on a computer screen, at eye-level, while playing on a keyboard at about waist-level). It is possible that a more realistic image of the hand in a more natural setting would have led to a stronger effect of visual feedback. One way to provide a more realistic image would have been to attach more markers to the hand. We tested such a solution and discovered that the improved image was accompanied by larger time lags in the streaming video, which we considered to be a more significant cost than the simplification of the hand image. Practical difficulties aside, there is reason to believe that we might find similar results even with a more realistic visual image. In the auditory domain, it is well known that the effects of DAF on speech remain when the speech signal is transformed to a non-speech signal (e.g., Howell and Archer 1984). Such results suggest that the effect of DAF is based on the timing of the auditory signal alone,

irrespective of whether the auditory signal is identifiably the result of produced actions. In a similar way, the disruptive effect of DVF may be based primarily on temporal interference within an abstract configuration that strongly suggests biological motion (cf. Johansson 1973).

Taken together, these results suggest that fluent sequence production relies on the appropriate timing of perceptual feedback across multiple modalities, here auditory and visual. The contribution of auditory and visual information to production is additive. Like previous research (Finney and Palmer 2003; Pfordresher 2005; Repp 1999), removing feedback did not disrupt production like delays did. Moreover, we did not find that the absence of feedback in one modality modulated the disruptive effect of delayed feedback from the other modality. We see the present results as consistent with an account of disruption from altered feedback that is still based on interference between perception and action (i.e., sensorimotor conflict) while highlighting the need to take into account the relative influences on production of feedback in all modalities.

Conflict of interest The authors declare that they have no conflict of interest.

References

Aschersleben G, Prinz W (1997) Delayed auditory feedback in synchronization. J Mot Behav 29:35–46

Dalla Bella S, Palmer C (2011) Rate effects on timing, key velocity, and finger kinematics in piano performance. PloS one 6:e20518

Finney SA (2001) FTAP: a Linux-based program for tapping and music experiments. Behav Res Meth Instrum Comput 33:65–72

Finney SA, Palmer C (2003) Auditory feedback and memory for music performance: sound evidence for an encoding effect. Mem Cognit 31:51–64

Flach R (2005) The transition from synchronization to continuation tapping. Hum Mov Sci 24:465–483

Gibson JJ (1966) The senses considered as perceptual systems. Houghton Mifflin, New York

Goebl W, Palmer C (2008) Tactile feedback and timing accuracy in piano performance. Exp Brain Res 186:471–479

Hommel B, Müsseler J, Aschersleben G, Prinz W (2001) The theory of event coding (TEC): a framework for perception and action planning. Behav Brain Sci 24:849–877

Hove MJ, Iversen JR, Zhang A, Repp BH (in press) Synchronizing with competing visual and auditory rhythms: bouncing ball meets metronome. Psychol Res

Howell P (2001) A model of timing interference to speech control in normal and altered listening conditions applied to the treatment of stuttering. In: Maassen B, Julstijn W, Kent R, Peters HFM, Van Lieshout PHMM (eds) Speech motor control in normal and disordered speech. Uttgeverij Vantilt, Nijmegen, pp 291–294

Howell P, Archer A (1984) Susceptibility to the effects of delayed auditory feedback. Percept Psychophys 36:296–302

Howell P, Sackin S (2002) Timing interference to speech in altered listening conditions. Acoust Soc Am 111:2842–2852

James W (1890) The principles of psychology, vol 2. Holt, New York

Johansson G (1973) Visual perception of biological motion and a model for its analysis. Percept Psychophys 14:201–211

Keller PE, Koch I (2008) Action planning in sequential skills: relations to music performance. Q J Exp Psychol 61:275–291

Kolers PA, Brewster JM (1985) Rhythms and responses. J Exp Psychol Hum Percept Perform 11:814–827

Langenberg U, Hefter H, Kessler K, Cooke J (1998) Sinusoidal forearm tracking with delayed visual feedback. I. Exp Brain Res 118:161–170

Lee DN, Reddish PE (1981) Plummeting gannets: a paradigm of ecological optics. Nature 293:293–294

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

Madison G, Merker B (2004) Human sensorimotor tracking of continuous subliminal deviations from isochrony. Neurosci Lett 370:69–73

Occelli V, Spence C, Zampini M (2011) Audiotactile interactions in temporal perception. Psychon Bull Rev 18:429–454

Patel AD, Iversen JR, Chen Y, Repp BH (2005) The influence of metricality and modality on synchronization with a beat. Exp Brain Res 163:226–238

Pfordresher PQ (2003) Auditory feedback in music performance: evidence for a dissociation of sequencing and timing. J Exp Psychol Hum Percept Perform 29:949–964

Pfordresher PQ (2005) Auditory feedback in music performance: the role of melodic structure and musical skill. J Exp Psychol Hum Percept Perform 31:1331–1345

Pfordresher PQ (2006) Coordination of perception and action in music performance. Adv Cognit Psychol 2:183–198

Pfordresher PQ (2008) Auditory feedback in music performance: the role of transition-based similarity. J Exp Psychol Hum Percept Perform 34:708–725

Pfordresher PQ, Dalla Bella S (2011) Delayed auditory feedback and movement. J Exp Psychol Hum Percept Perform 37:566–579

Pfordresher PQ, Kulpa JD (2011) The dynamics of disruption from altered auditory feedback: further evidence for a dissociation of sequencing and timing. J Exp Psychol Hum Percept Perform 37:949–967

Pfordresher PQ, Palmer C (2002) Effects of delayed auditory feedback on timing of music performance. Psychol Res 66:71–79

Prinz W, Aschersleben G, Koch I (2009) Cognition and action. In: Morsella E, Bargh J, Gollwitzer PM (eds) Oxford handbook of human action. Oxford University Press, Oxford, pp 35–71

Repp BH (1999) Effects of auditory feedback deprivation on expressive piano performance. Music Percept 16:409–438

Repp BH, Penel A (2002) Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. J Exp Psychol Hum Percept Perform 28:1085–1099

Repp BH, Penel A (2004) Rhythmic movement is attracted more strongly to auditory than to visual rhythms. Psychol Res 68:252–270

Smith WM, McCrary JW, Smith KU (1960) Delayed visual feedback and behavior. Science 132:1013–1014

van den Heuvel MRC, Balasubramaniam R, Daffertshofer A, Longtin A, Beek PJ (2009) Delayed visual feedback reveals distinct time scales in balance control. Neurosci Lett 452:37–41

Yates AJ (1963) Delayed auditory feedback. Psychol Bull 60:213–232