Journal of Experimental Psychology:
Human Perception and Performance
2016, Vol. 42, No. 5, 658–670

0096-1523/16/$12.00 http://dx.doi.org/10.1037/xhp0000172

Exploring Perception–Action Relations in Music Production: The Asymmetric Effect of Tonal Class

Andrew T. Jebb
University at Buffalo, State University of New York
and Purdue University

Peter Q. Pfordresher
University at Buffalo, State University of New York

When playing musical passages, performers integrate the pitch content of auditory feedback with current action plans. However, this process depends on the degree to which the musical structure of the feedback melody is perceived as similar to the structure of what is planned. Four experiments reported here explored the relationship between the *tonal class* of planned melodies (tonal or atonal) and the sequence of events formed by auditory feedback. Participants produced short melodies from memory that were either tonal (Experiments 1 and 3) or atonal (Experiments 2 and 4). Auditory feedback matched the planned melody with respect to contour but could vary in tonal class. The results showed that when participants planned a tonal melody, atonal feedback was treated as unrelated to the planned sequence. However, when planning an atonal melody, tonal feedback was still treated as similar to the planned sequence. This asymmetric similarity mirrors findings found within the music perception literature and implies that schematic musical knowledge is highly active in determining perception–action relations during music performance.

Keywords: auditory feedback, sequence production, music cognition, perception and action, tonality

In a wide range of behaviors, humans rely on and exploit the perceptual information that results from the performance of actions, called *perceptual feedback*. Although perceptual information is behaviorally valuable in general, feedback is unique in its relevance for *concurrent* action planning and coordination. Imagine singing or maneuvering an automobile barred from perceiving the effects of your produced actions: The consequences may be embarrassing or even dangerous. Further, understanding the role of perceptual feedback in the planning and execution of action sequences offers valuable insights into the nature of sensorimotor integration.

In this article, we focus on the role of *auditory feedback* in music performance and investigate how structural features of music influence perception–action relationships. Auditory feedback constitutes the most important source of information about performance accuracy, given that a particular pattern of sounds is the goal of every performance. Further, music, like speech, involves the organization of basic units into long, complex sequences based on grammatical rules (Lerdahl & Jackendoff, 1983). Although musical sequences are composed of single events, it is their *interrelationships* that create the types of musical structure that dominate music perception. This structure includes surface features, such as how the pitches rise and fall across the melody (melodic contour) and the magnitude of change across pitches (pitch interval), as well as deeper grammatical features, such as whether the music contains a tonal center (tonal class) and what that tonal center is (i.e., the key).

In music production, two sequences are continuously integrated in real time: the *planned* melody and the *feedback* melody. The latter simply refers to the perceived pattern of auditory events, whereas the planned melody comprises both the series of actions (e.g., key presses on a piano) and their expected outcomes. Under ordinary performance conditions, these two sequences are matched in both their *timing* (i.e., the synchronization of events) and *content* (event categories associated with sound, such as a G# note or C major chord). In other words, correctly performing actions produces feedback that is consistent with the expected musical event. However, manipulating feedback can provide insights into how feedback interacts with current action plans. The altered auditory feedback (AAF) paradigm has been used for over half a century to explore perception–action associations and consists of participants producing a sequence of actions (speech, music, Morse code, etc.) while hearing auditory feedback whose relationship to these actions is altered. The first AAF paradigms involved delaying the onsets of auditory feedback, leading to asynchronies of perception and action (for reviews, see Finney, 1999; Howell, 2004; Yates, 1963). Such timing manipulations generally manifest in timing errors in musical performance (e.g., slower production) rather than

---

Andrew T. Jebb, Department of Psychology, University at Buffalo, the State University of New York, and Department of Psychological Sciences, Purdue University; Peter Q. Pfordresher, Department of Psychology, University at Buffalo, the State University of New York.

This research was sponsored by National Science Foundation Grants BCS-0642592 and BCS-1256964. This work represents a portion of a master's thesis completed by the first author at the University at Buffalo. We thank Stefanie Acevedo for programming the Max patches and collecting pilot data for Experiment 1, Michael Williams for assisting with data collection, and Matthew McCabe and Bruno Repp for helpful advice on programming altered feedback designs in Max.

Correspondence concerning this article should be addressed to Peter Q. Pfordresher, University at Buffalo, 362 Park Hall, Buffalo, NY 14260. E-mail: pqp@buffalo.edu

mistakes in sequencing (i.e., "wrong notes"; Pfordresher, 2006). By contrast, altering the content of feedback tends to produce sequencing errors, not timing errors, and the effects of these manipulations are comparatively less understood. In other words, it is yet unclear how different structural features of music influence perception–action relationships.

The Role of Melodic Contour and Tonal Class

The experiments reported here use the AAF paradigm to investigate the role of musical content in sensorimotor integration. More specifically, we explored how the *tonal class* of musical sequences determines perception–action relations in music production. Previous work has shown that, for altered feedback content to cause disruption, the planned and feedback melodies need to be perceived with some degree of similarity in musical structure. For instance, when AAF changes the normal feedback to a sequence of random pitches, little or no disruption is caused because the two melodies have no *structural similarity* (Finney, 1997; Pfordresher, 2005). However, when the feedback is manipulated in such a way that its structure remains similar to the planned melody, performance may be disrupted. One example of such a manipulation is a *serial shift*, in which every event produced on the instrument leads to a feedback event normally associated with a different position in the planned sequence. For example, a lag-1 serial shift causes every key press on a piano to produce the pitch normally associated with the previous key press. Here, the order of feedback pitches is shifted behind (i.e., "lags") the planned events by one serial position. Theoretically, serial shifts cause disruption because perception and action share a common representation of sequence structure (Pfordresher, 2006; Pfordresher & Kulpa, 2011), and hearing feedback associated with different actions activates action plans intended for other positions in the melody.

Using serial shifts to investigate how musical structure influences perception–action relations in music, Pfordresher (2008) demonstrated that disruption occurs when the feedback melody shares the same *melodic contour* as the planned melody (i.e., the pattern of upward or downward motion). In other words, similarity, and thus disruption, were present when contour was preserved, even though other structural features, such as the musical key and individual pitch intervals, were altered. Thus, contour preserves the structural similarity between the planned and feedback sequences even when no individual pitches are held in common.

Another important structural feature of music is referred to as *tonality* and is the topic of the present investigation. Tonality refers to a schema that hierarchically organizes each pitch within an established key. Western music allows 12 possible pitch classes (*chromas*), each of which can serve as the tonal center around which the other pitches vary in their relative importance in the hierarchy (e.g., in the key of "C," the pitches C, E, and G are most important). Music that is organized in this way is called *tonal*, and tonal melodies dominate Western music; tonal relationships are a salient feature of melodies for listeners regardless of training (Bigand & Poulin-Charronnat, 2006; Krumhansl, 1990; Krumhansl & Cuddy, 2010). Tonal relationships also play a role in performance, influencing the kind of pitch errors trained pianists make (Palmer & van de Sande, 1993) and the visual processing of movements on a keyboard (Novembre & Keller, 2011).

In the present research, we consider a distinction between melodies that are tonal and those that are *atonal*, a distinction we refer to as *tonal class.* Atonal music, which became prominent in the early 20th century, dispensed with the standard hierarchical organization of tonal music and freely used all 12 pitch classes. Most relevant to the current study, atonal melodies eschew the grammatical structure of tonal music, so the role of structural similarity regulating perception and action may function differently for these sequences than for tonal melodies.

Previous studies suggest that tonal class, along with melodic contour, constrains associations between perception and action. In the aforementioned study by Pfordresher (2008), participants performed tonal melodies from memory. When the feedback melody was altered to be an *atonal* variation, serial shifts did not disrupt production even though melodic contour was preserved. Thus, changing tonal class eliminated the perceived structural similarity based on contour. Unfortunately, a major limitation of Pfordresher (2008) was that it only examined the effect of tonal class on planned tonal melodies and did not examine the kind of sequential perception–action associations that are formed when planning atonal sequences. The role of tonal class is particularly critical because it bears on the underlying grammatical structure of musical sequences and suggests perception–action relationships are also based on schematic knowledge.

The Present Experiments

In this study, four experiments were designed to investigate tonal class as a structural factor in determining perception–action similarity. Participants memorized and performed tonal melodies (Experiments 1 and 3) or atonal melodies (Experiments 2 and 4). Feedback melodies always matched the contour of the planned melody but could differ (in AAF conditions) in tonal class and also be serially shifted. The underlying logic of these manipulations was that serial shifts would only be disruptive when the feedback melody was structurally similar to the planned melody. If serially shifting the feedback melody failed to produce disruption when its tonal class is changed, then the performer is likely treating the feedback melody as irrelevant to the planned sequence.

We considered three possible outcomes of these manipulations, each having a distinct theoretical implication. Hypothesis 1 proposed that tonal and atonal melodies are not at all structurally similar. Thus, when participants perform atonal melodies, serial shifts of tonal variations should fail to disrupt, implying that different tonal classes are nonoverlapping with respect to structural similarity.

Hypothesis 2 followed from the prediction that similarity relationships across tonal classes may be *asymmetric* (cf. Tversky, 1977). As described earlier, tonal music is hierarchically organized and commands strong expectations about what the future pitches will be. These schema-driven expectations may cause any atonal feedback to be perceived as structurally unrelated within music performance. Conversely, because they lack this salient organization, atonal melodies may not produce strong expectations and may therefore fail to produce the same experience of dissimilarity when tonal feedback is heard. If so, then serial shifts of tonal variations may disrupt the production of a planned atonal melody. Some support for this hypothesis comes from the music perception literature. When tonal melodies are presented prior to atonal melodies, a greater sense of dissimilarity is reported than when the presentation order is reversed called "asymmetric similarity" (i.e., than when atonal melodies are followed by tonal sequences; Bartlett & Dowling, 1988; Dowling & Bartlett, 1981).

Finally, we considered a third hypothesis based on the common subjective experience of atonal melodies as "formless." It may be the case that a planned atonal melody leads to no expected pitch outcomes whatsoever and that performers simply plan these melodies as a sequence of motor movements. If so, then experiments with planned atonal melodies should lead to no effects of AAF whatsoever.

Although music performance is typically considered a specialized skill, our focus is primarily on perception–action associations in the general population. Our use of a music performance task is based primarily to explore the role of sequential grammars, rather than to understand the effects of skill acquisition. Past research suggests similar qualitative effects of AAF for musicians and nonmusicians, although the magnitude of disruption may differ in degree (Pfordresher, 2005, 2012). In this context, we are primarily interested in the role of tonal class as an implicitly learned grammar as opposed to explicit categorization responses that result from extensive training (Tillmann, Bharucha, & Bigand, 2000).

Experiment 1

Experiment 1 was a conceptual replication and extension of Pfordresher (2008). Participants performed previously unfamiliar tonal melodies on a keyboard while listening to auditory feedback over headphones. The auditory feedback sequence was either a transposed version of the planned tonal melody (cf. Pfordresher, 2008, Experiment 1) or an atonal sequence that maintained its contour but shared no pitches or pitch intervals with the planned melody. In addition, either feedback melody could also have a lag-1 serial shift. We predicted the results would replicate Pfordresher (2008): Serial shifts of the tonal melody would disrupt accuracy (measured as error rates per trial), whereas the shifted and nonshifted atonal variations would not disrupt performance.

Method

Participants

Twenty-four undergraduate students (12 men, 12 women) participated in the present study in exchange for credit toward an introductory psychology course. There were no prerequisites related to musical experience, though subjects reported $M = 1.75$ years of piano training (range = 0–10) and $M = 2.79$ years of nonpiano training (range = 0–9). Eleven participants (46% of the sample) reported at least one year of piano training ($M = 3.8$, max = 10), but only four of these (16% of the entire sample) had more than five years and could be considered "pianists." All participants reported being right-handed, and no participant reported hearing deficits or any impairment in the use of their dominant hand. None reported having absolute pitch. The mean age across participants was 19.63 years (range = 18–30).

Materials

Four melodies were created for the present study, each consisting of eight notes and a distinct melodic contour. Melodies were monophonic and played with the right hand. Five pitch classes were chosen for each melody and were the first five pitch classes of the major scale in the musical key of the tonal melodies. The melodies were restricted so that each finger could be assigned to one of five adjacent white keys, making production of the melodies straightforward for musically untrained participants. No melody included repeating pitches on successive events.

The music notation for learning and memorization of the melodies used numbers 1–5 to represent finger-key pairings (e.g., 1 = thumb), with these numbers also positioned above each respective key on the keyboard. On the notation, a row of numbers corresponding to melody pitches (finger-key mappings) was displayed, and above each number there was a small diagram of a hand with the appropriate finger highlighted (e.g., see Pfordresher, 2005; Pfordresher & Beasley, 2014). Each melody was to be played in binary meter (4/4 time) with each note played as a quarter note. This was communicated by a vertical line separating the first four notes of the melody from the last four and verbal instructions to the participant (see Procedure).

Conditions

Participants memorized a keyboard fingering sequence that was associated with a tonal melody. In each experimental trial, auditory feedback was set to one of four conditions: tonal feedback (normal feedback), serially shifted tonal feedback, atonal feedback, or serially shifted atonal feedback. Figure 1 shows the relationship between a planned melody and the four types of feedback melodies. For those conditions in which tonal class was altered (i.e., the two conditions with atonal feedback), the pitches of the original tonal melody were mapped onto an atonal, contour-preserving variant using the software program *Max 6* (Cycling '74). The five adjacent white keys used during performance (MIDI numbers 60–67) produced chromatic pitches (C4# to F4) as opposed to their standard pitches in the key of C major. In conditions with tonal feedback (feedback melodies on the left in Figure 1), these five keys also produced altered pitches, the first five pitches of the key of C# major. This was done to ensure that the tonal conditions were not advantaged in being consistent with the original key–pitch associations of the keyboard. The lag-1 serial shift manipulation (lower row of feedback melodies in Figure 1) was also produced by the *Max* software program and involved a function that mapped the input key to the output pitch of the previous key.

We confirmed that the atonal melodies were indeed atonal (i.e., not implying any particular key) using the Krumhansl-Schmuckler algorithm (Krumhansl, 1990, pp. 78–81). Because of the short length of melodies, we aggregated pitch classes across all tonal melodies and constructed a frequency distribution of all pitch classes. This frequency distribution was then correlated with a set of 24 major and minor keys, the strongest correlation indicating the most likely key. As expected, the most likely key for the aggregated tonal feedback melodies was C# major, $r(10) = .72$, $p < .01$. By contrast, atonal melodies yielded a near-zero correlation with C# major, the most likely key based on finger positioning and associations with the planned melody, $r(10) = .06$. The atonal feedback melodies correlated most strongly with the key of D minor, but this relationship was not significant and much weaker than that observed for the tonal melodies, $r(10) = .39$, $p = .50$.

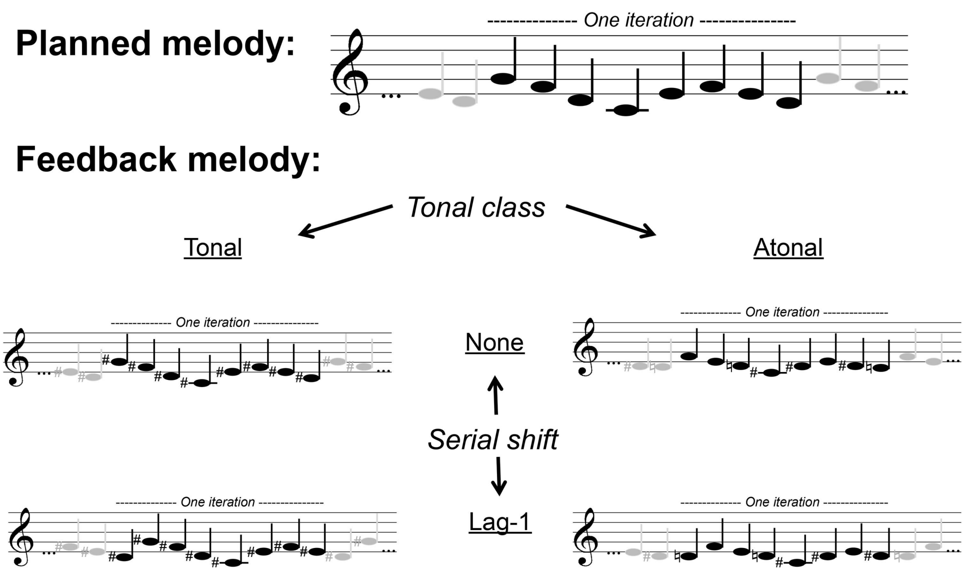

Figure 1. Examples of stimulus conditions for one melody, shown using music notation that excludes meter bars and key signatures for clarity. Because sequences were performed repeatedly, notation highlights the pattern of produced and feedback events of one 8-event iteration, preceded and followed by other iterations (implied by use of greyed notes and ellipses). Notation for the planned melody indicates the piano keys that the participant intends to press when not making errors. Notation for feedback melodies indicates the sounded pitches that would accompany produced keys.

Apparatus

Participants performed melodies on an M-Audio Keystation 49e keyboard, which was set to a height that was anticipated to be comfortable but adjusted if subjects expressed discomfort when asked. Auditory feedback was heard through Sony MDR-7506 Professional headphones set to a comfortable level. All feedback was produced by the *Max* software package, which also recorded all performance data.

Procedure

Participants were seated before a keyboard and fitted with headphones. One melody had been placed in front of the keyboard prior to arrival, and an explanation of the notation was given. Participants demonstrated that they understood the notation by playing the melody correctly several times with the notation present. Participants were then informed that the notation would not be made available during the experimental trials and then proceeded to freely learn the melody for as long as they considered necessary for memorization. Participants notified the experimenter when they had memorized the melody and then demonstrated this by performing it correctly three times in succession with the notation removed.

After memorization, participants were given a perceptual test to verify that they had internalized the sound of the melody in addition to the motor movements. The perceptual test consisted of playing two melodies in succession over the participants' headphones: the learned melody and a similar-sounding melodic lure that had been altered by 1–3 pitches (the number varied in order to maintain an approximately equal level of difficulty across lures). After hearing both melodies, participants were asked to identify the melody they had just learned. This test was administered twice using two different lures per melody. If either answer was incorrect, participants were asked to spend further time learning the melody. Actual initiation of the experiment occurred only after (a) both initial answers were correct or (b) participants gave at least one correct answer after taking additional time learning the melody (which occasionally required more than one repetition).

Participants then performed a practice trial. The musical notation was taken away, and subjects were instructed to begin playing after eight isochronous beats of a metronome that communicated the target tempo [500-ms interonset intervals (IOIs), or 120 beats per minute]. Further instructions were given to ignore any mistakes made (i.e., to not try and correct errors) and to repeatedly play the melody to the best of their ability until the end of the trial. Here, auditory feedback was unaltered (i.e., the tonal, no serial shift condition), and this trial lasted 80 key presses, at which point a quiet bell sound signaled its conclusion.

The experimental trials immediately followed the practice trial and were identical in form, the only difference being the introduction of AAF. Each participant performed two blocks, each of 16 trials, and played one melody per block. This practice yielded a total of 32 experimental trials. In every four-trial cycle, each feedback condition was used once. Within each four-trial cycle, the order of conditions was varied to counterbalance the feedback alterations.

After each trial, participants rated its difficulty on a 7-point Likert scale, ranging from 1 (*least difficult*) to 7 (*most difficult*). Participants completed a questionnaire pertaining to musical and language background after the first block.

Analysis

The two dependent measures of interest in this study were error rates and difficulty ratings, both of which relate to the effect of AAF manipulations on planning. Error rates per trial (number of pitch errors relative to 80 total keypresses) were based on a computer algorithm (Large, 1993; Palmer & van de Sande, 1993, 1995) that computed the number of modifications (insertions, deletions, substitutions) needed to make the produced sequence match the target sequence (cf. Levenshtein, 1966).

Error rates and difficulty ratings were submitted to a 2 (tonal class: tonal, atonal) × 2 (serial shift: shift, no shift) repeated-measures analysis of variance (ANOVA). Measures of effect size included partial eta squared (η^2_{p}) and generalized eta squared (η^2_{G}; Olejnik & Algina, 2003).[1] In addition, we ran planned comparisons (paired *t* tests) between each experimental AAF condition and the control (normal feedback) condition. Each test adopted the one-tailed prediction that error rates and difficulty ratings should increase with AAF. Following recommendations of Keppel (1991, p. 169; see also Klockars & Sax, 1986, p. 38) for the modified Bonferroni correction, we retained a testwise critical alpha level of .05. Measures of effect size for planned comparisons were based on Cohen's *d* (Cohen, 1988).

In addition to error rates and difficulty ratings, we also analyzed performance timing. Timing measures were of less interest because they are less related to the planning of serial order and typically yield null effects for the types of AAF content manipulations used here. However, there were more timing effects in the present experiments than in other studies. For each trial, we analyzed the mean of produced inter-onset intervals (IOIs) and the coefficient of variation ($CV = SD/M$) for IOIs after removing outliers (IOIs outside 2 *SD*s around the mean within a trial) and any linear trend between IOI and sequence position (e.g., a gradual slowing across the course of a trial). These measures were analyzed in the same way as error rates and difficulty ratings.

Results

The two-way repeated-measures ANOVA on error rates revealed a main effect of serial shift, $F(1, 23) = 6.18$, $p = .021$, $\eta^2_{\text{p}} = .21$, $\eta^2_{\text{G}} = .033$, no main effect of tonal class ($p = .19$, $\eta^2_{\text{p}} = .067$, $\eta^2_{\text{G}} = .002$), and no interaction effect ($p = .67$, $\eta^2_{\text{p}} = .009$, $\eta^2_{\text{G}} < .001$; see Figure 2A). Though an interaction effect was not found in the omnibus ANOVA, planned comparisons effectively illustrated such a pattern. In accordance with past research, serially shifting the planned tonal sequence ($M = 0.067$, $SD = 0.055$) resulted in significant disruption relative to normal feedback ($M = 0.049$, $SD = 0.04$), $t(23) = 1.83$, $p = .04$, $d = 0.39$, but serially shifting the atonal feedback condition ($M = 0.061$, $SD = 0.047$) did not reach statistical significance ($p = .07$, $d = 0.28$). Nonshifted atonal feedback ($M = 0.047$, $SD = 0.039$) did not increase error rates either ($p = .33$, $d = -0.06$), suggesting that all atonal feedback was treated as irrelevant to the planned melody.

The analysis of subjective difficulty ratings (Figure 3A) yielded different results: a main effect of serial shift, $F(1, 23) = 22.33$, $p < .001$, $\eta^2_{\text{p}} = .49$, $\eta^2_{\text{G}} = .11$, a main effect of tonal class, $F(1, 23) = 9.04$ $p = .006$, $\eta^2_{\text{p}} = .28$, $\eta^2_{\text{G}} = .04$, and a significant interaction, $F(1, 23) = 12.71$, $p = .002$, $\eta^2_{\text{p}} = .36$, $\eta^2_{\text{G}} = .03$. Planned comparisons verified that subjects reported all altered feedback conditions as more difficult than the normal feedback condition ($M = 2.51$, $SD = 1.17$). Difficulty ratings increased for the nonshifted atonal condition ($M = 3.32$, $SD = 1.11$), $t(23) = 4.07$, $p < .001$, $d = 0.71$, the serially shifted atonal condition ($M = 3.66$, $SD = 1.06$), $t(23) = 4.94$, $p < .001$, $d = 1.03$, and the serially shifted tonal condition ($M = 3.63$, $SD = 1.04$), $t(23) = 4.80$, $p < .001$, $d = 1.01$. These large effect sizes (Cohen, 1988) indicate that participants perceived each altered feedback condition as substantially more difficult than performing with normal feedback, despite the fact that only the shifted tonal condition significantly increased errors in performance.

Measures of performance timing (mean IOI and CV of IOI) from Experiment 1 are shown in Table 1. Consistent with prior research using similar manipulations, effects of feedback condition on timing measures were not significant ($p > .10$ for all main effects and interactions). Likewise, no planned comparisons were significant for differences in mean IOI. However, for the CV of IOI, planned comparisons between the control condition ($M = 0.17$, $SD = 0.12$) and serially shifted feedback conditions were significant: shifted tonal ($M = 0.19$, $SD = 0.13$), $t(23) = 1.84$, $p = .04$, $d = 0.17$, shifted atonal ($M = 0.19$, $SD = 0.13$), $t(23) = 2.13$, $p = .02$, $d = 0.12$. It is worth noting that the dissociation between sequencing and timing found in similar research is strongest when contrasting the effects of altered feedback on error rates versus mean IOI. By contrast, when timing variability is measured (CV of IOI), effects of altered feedback are more similar to effects found on error rates (Pfordresher, 2003).

Discussion

Experiment 1 replicated the findings of Pfordresher (2008) such that when planning a tonal melody, hearing a serially shifted atonal variation did not cause disruption. Even though melodic contour was preserved across these sequences, the change in tonal class voided their structural similarity. It is interesting that this pattern emerged even as the primary dependent error measures differed across these two studies. In Pfordresher (2008), errors were defined as the proportion of trials containing at least one error, here, the proportion of errors within a trial. Therefore, across two different metrics, tonal class determined how performers related auditory feedback to cognitive action planning. The lack of significant disruption suggests that hearing atonal or serially shifted atonal feedback failed to activate actions meant for alternate serial positions. Subjects treated this feedback as irrelevant to the planning and coordination of future actions in a manner analogous to hearing random feedback (Finney, 1997; Pfordresher, 2005).

Unlike Pfordresher (2008), we also analyzed subjective difficulty ratings, and these results differed in important ways from error rates. Subjects reported all trials with altered auditory feedback as more difficult than the control condition (with large effect sizes), despite only the serially shifted tonal condition actually resulting in greater performance errors. The most plausible explanation is that subjects were genuine in their

---

[1] Although partial eta square is a more common measure, we also include generalized eta square because it allows for effect size comparisons across research designs (e.g., repeated measures, between-subjects, designs with nested, blocked or random factors; Bakeman, 2005; Fritz, Morris, & Richler, 2012). Note that generalized eta square yields a smaller number because it includes estimates of nonmanipulated variance sources in the denominator.

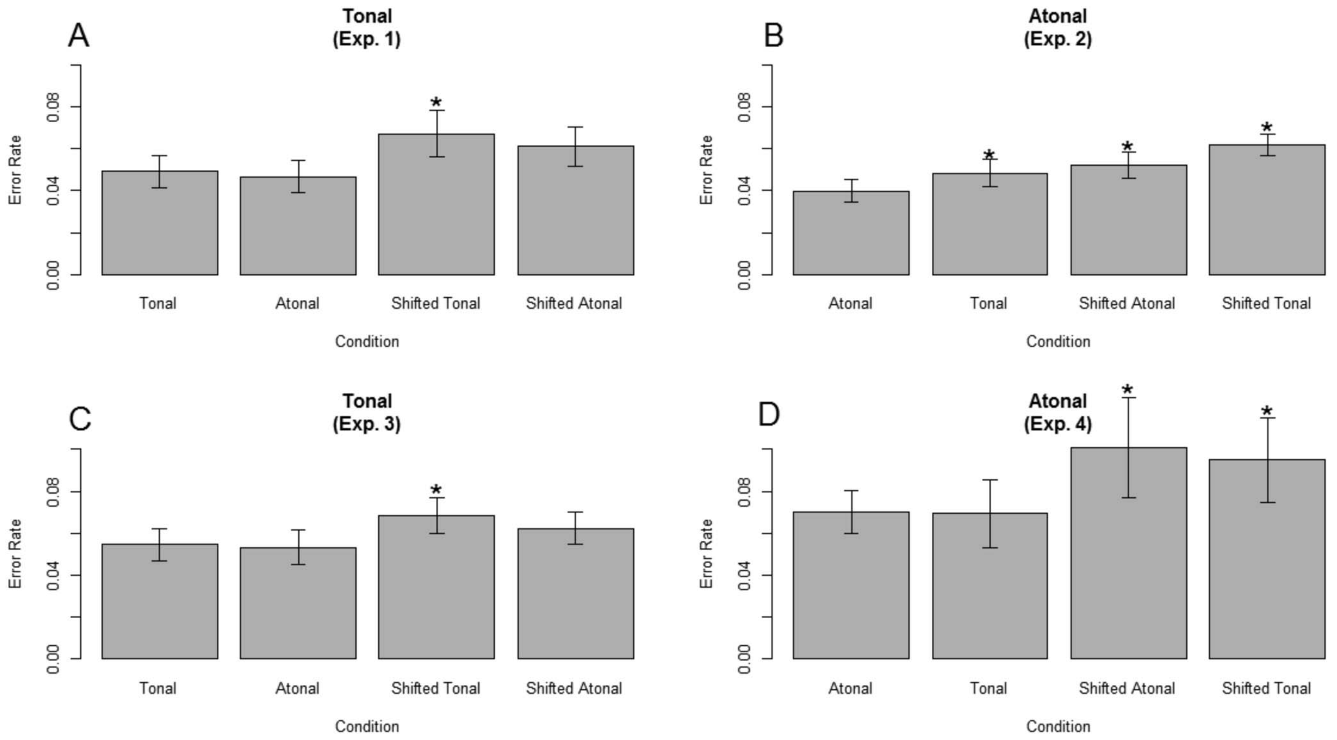

Figure 2. Mean error rates of each condition across experiments (Exp.). Comparisons were made against the normal unaltered feedback condition, shown at the far left of each panel, and stars represent a significant difference at $p < .05$. Error bars represent 1 *SE* of the mean.

reporting and found it difficult to ignore the strange sounding altered feedback. However, error rates did not increase as no alternate action plans were systematically activated. This result converges with other recent evidence suggesting that conscious error monitoring and disruption of production from AAF may involve separate resources (Pfordresher & Beasley, 2014).

Experiment 2

Experiment 1 found that hearing feedback melodies differing in tonal class caused subjects to treat this perceptual information as dissimilar to the planned melody. However, like Pfordresher (2008), Experiment 1 was limited in that the planned melody was always tonal, and changes in tonal class were always changes from tonal to atonal melodies. Thus, Experiment 2 complements Experiment 1 by testing whether tonal feedback melodies are treated as dissimilar to a planned atonal melody.

One challenge in performing atonal melodies on a standard keyboard is that such pitch intervals typically require using both the white and black keys, which significantly increases the motor challenge of performance based on ergonomic considerations (Parncutt, Sloboda, Clarke, Raekallio, & Desain, 1997; Sloboda, Clarke, Parncutt, & Raekallio, 1998). As such, we devised a setup in which the same finger patterns used in Experiment 1 led to an atonal feedback sequence that would be associated with actions during learning. This was done by matching piano keys to a fixed set of alternate pitches in *Max*. However, this created an alternate mapping between keys and associated pitches that involved a more salient departure from the standard mapping of the piano. Whereas the altered mapping in Experiment 1 only involved a one-semitone shift of all keys (from C major to C# major), the pitch range in Experiment 2 was significantly compressed. Though this could be problematic, we anticipated that participants would have little difficulty learning melodies with the altered mapping, as other research suggests that nonpianists can form pitch mapping rapidly (Bangert & Altenmüller, 2003; Lahav, Saltzman, & Schlaug, 2007). The issue of pitch range was more fully addressed in Experiments 3–4.

Method

Participants

Twenty-four undergraduate students (11 men, 13 women) participated in exchange for course credit. Participants reported M = 0.88 years of piano training (range = 0–4) and M = 1.79 years of nonpiano training (range = 0–11). As in Experiment 1, 11 participants (46% of the sample) reported one or more years of piano lessons; however, none of these participants reported more than 5 years of lessons (M = 1.9, max = 4). One participant reported being left-handed, and no participant reported hearing deficits or any impairment in the use of their dominant hand. One reported having absolute pitch. Their mean age was 18.75 years (range = 18–25).

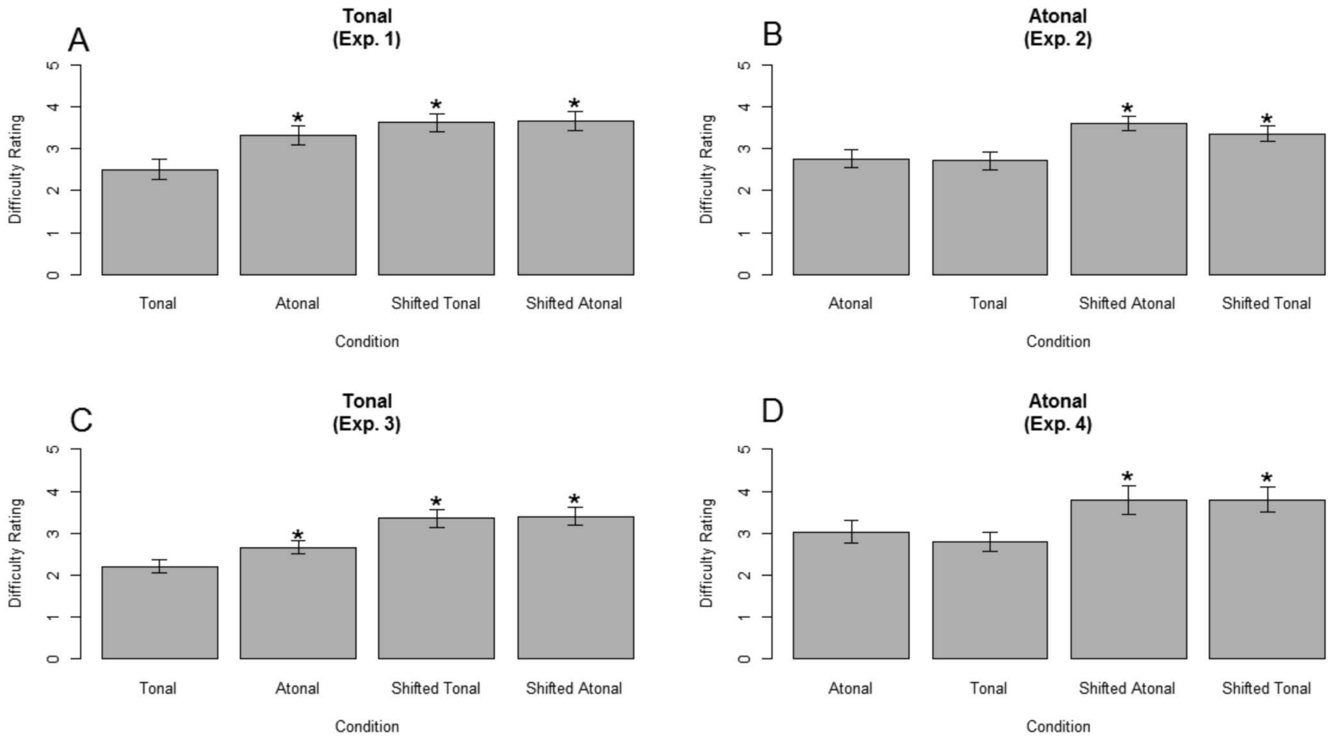

Figure 3. Mean difficulty ratings of each condition across experiments (Exp.). Comparisons were made against the normal unaltered feedback condition, shown at the far left of each panel, and stars represent a significant difference at $p < .05$. Error bars represent 1 *SE* of the mean.

Conditions, Materials, Procedure, and Analysis

The conditions, materials, and analysis of Experiment 2 were identical to Experiment 1. The only difference was the tonal class of the learned melody. Each participant still learned two of the four melodies, but the learned melody (i.e., normal feedback) was changed to be the atonal variant of Experiment 1. Thus, participants learned and planned an atonal melody, and during the experiment, they were exposed to four feedback conditions: normal (now atonal), serially shifted atonal feedback, a tonal variant, and this tonal variant serially shifted.

Results

A two-way repeated-measures ANOVA on error rates (Figure 2B) resulted in a main effect of serial shift, $F(1, 23) = 9.31$, $p =$

Table 1
Tempo (Mean IOI) and Timing Precision (CV IOI) From Experiments 1–4 Averaged Across Participants and Trials

	Serial shift			
Experiment	Tonal class (nonshifted conditions)		Tonal class (shifted conditions)	
1	Tonal	Atonal	Tonal shifted	Atonal shifted
Mean IOI	460.15 (14.31)	462.31 (13.15)	468.43 (12.18)	461.19 (12.40)
CV IOI	.17 (.026)	.18 (.030)	**.19** (.026)	**.19** (.027)
2	Atonal	Tonal	Atonal shifted	Tonal shifted
Mean IOI	484.32 (16.36)	481.49 (17.32)	**500.06** (19.02)	**497.28** (19.11)
CV IOI	.26 (.038)	.24 (.034)	**.28** (.035)	**.28** (.036)
3	Tonal	Atonal	Tonal shifted	Atonal shifted
Mean IOI	417.92 (16.29)	412.17 (16.64)	**432.23** (18.35)	**428.71** (17.52)
CV IOI	.24 (.028)	.23 (.028)	**.26** (.024)	.26 (.025)
4	Atonal	Tonal	Atonal shifted	Tonal shifted
Mean IOI	466.84 (23.42)	462.34 (22.30)	486.58 (26.43)	**484.19** (25.74)
CV IOI	.254 (.025)	.257 (.025)	**.287** (.028)	**.281** (.027)

Note. Parentheses report standard errors. Boldface text shows significant differences from the control condition for each experiment (shown in the left-most column). IOI = interonset intervals; CV = coefficient of variation (SD IOI/M IOI).

$.006$, $\eta^2_p = .29$, $\eta^2_G = .05$, a significant main effect of tonal class, $F(1, 23) = 6.03$, $p = .022$, $\eta^2_p = .21$, $\eta^2_G = .03$, and no interaction effect ($p = .87$, $\eta^2_p = .001$, $\eta^2_G < .001$). In this respect, results were identical to Experiment 1. However, planned comparisons showed that, in contrast to Experiment 1, all altered feedback conditions resulted in more errors relative to the normal feedback condition ($M = 0.040$, $SE = 0.006$). A serial shift of the planned feedback ($M = 0.052$, $SE = 0.006$) led to an increase in error rates, $t(23) = 2.54$, $p = .009$, $d = 0.43$, arguing against Hypothesis 3, that serially shifting an atonal melody would not result in more errors. Moreover, the nonshifted tonal condition ($M = 0.048$, $SE = 0.006$) increased errors, $t(23) = 2.14$, $p = .04$, $d = 0.29$, as did serially shifting this tonal feedback ($M = 0.062$, $SE = 0.005$), $t(23) = 3.74$, $p < .001$, $d = 0.83$, arguing against Hypothesis 1 (viz., that including both a serial shift and change of tonal class would not result in greater errors). Effect sizes ranged from small to large (Cohen, 1988), though serially shifting the tonal feedback led to the largest effect ($d = 0.83$), relative to the serially shifted atonal ($d = 0.43$) and nonshifted tonal conditions ($d = 0.29$). It is worth noting that error rates in the control condition from Experiment 2 were slightly (though not significantly) lower than in Experiment 1 ($M = 0.040$ vs. 0.049, respectively), thus arguing against the possibility that the unusual key-to-pitch mapping in Experiment 2 made the experiment more difficult in general than Experiment 1.

As for the difficulty ratings (Figure 3B), a repeated-measures ANOVA yielded a main effect of serial shift, $F(1, 23) = 15.93$, $p = .001$, $\eta^2_p = .41$, $\eta^2_G = .13$, no main effect of tonal class, ($p = .065$, $\eta^2_p = .14$, $\eta^2_G = .01$), and no interaction ($p = .11$, $\eta^2_p = .11$, $\eta^2_G = .002$). Both the shifted atonal condition ($M = 3.59$, $SE = 0.17$) and the shifted tonal condition ($M = 3.36$, $SE = 0.18$) were rated as more difficult than normal feedback: $t(23) = 4.11$, $p < .001$, $d = 0.87$, and $t(23) = 3.09$, $p = .005$, $d = 0.62$, respectively. The nonshifted tonal condition ($M = 2.71$, $SE = 0.22$) was actually rated as easier than normal feedback ($M = 2.76$, $SE = 0.22$), but the difference between means did not reach statistical significance, $t(23) = 0.58$, $p = .57$, $d = -0.04$.[2]

As in Experiment 1, we also analyzed the mean and variability (CV) of produced IOIs during each trial. Descriptive statistics are shown in Table 1. In contrast to Experiment 1 and other similar studies, feedback manipulations influenced performance timing. For mean IOI, the ANOVA yielded a significant main effect for shift, $F(1, 23) = 13.71$, $p = .001$, $\eta^2_p = .37$, $\eta^2_G = .01$, but no main effect of tonal class ($p = .56$, $\eta^2_p = .02$, $\eta^2_G < .001$), and no interaction ($p = .99$, $\eta^2_p < .01$, $\eta^2_G < .001$). Planned comparisons showed significant differences between the nonshifted atonal condition ($M = 484.32$, $SD = 80.13$) and both serially shifted conditions: the shifted atonal melody ($M = 500.06$, $SD = 93.18$), $t(23) = 2.61$, $p = .008$, $d = 0.18$, and the shifted tonal variation ($M = 497.28$, $SD = 93.64$), $t(23) = 1.70$, $p = .05$, $d = 0.15$. However, the contrast with the nonshifted tonal variation was not significant ($M = 481.49$, $SD = 84.85$), $t(23) = -0.54$, $p = .70$, $d = -0.03$. These results qualitatively mirror the effects of altered feedback on error rates. However, subtle but important differences between error rates and mean IOI exist. First, whereas serial shifts of the tonal variation had a larger effect on error rates than did shifts of the atonal variation ($d = 0.83$ vs. 0.43, respectively), serial shifts of the tonal variation just barely crossed the significance threshold for mean IOI and yielded a smaller effect size than did serial shifts of the planned melody. Furthermore, whereas the nonshifted tonal variation elevated error rates, it had no effect on mean IOI.

Similar to mean IOI, the ANOVA on timing variability (CV IOI) yielded a significant main effect for shift, $F(1, 23) = 10.65$, $p = .003$, $\eta^2_p = .32$, $\eta^2_G = .01$, but not for tonal class ($p = .26$, $\eta^2_p = .06$, $\eta^2_G < .001$) or the interaction ($p = .40$, $\eta^2_p = .03$, $\eta^2_G < .001$). Planned comparisons yielded significant contrasts between the nonshifted atonal condition ($M = 0.26$, $SD = 0.19$), as well as the shifted atonal melody ($M = 0.28$, $SD = 0.17$), $t(23) = 2.16$, $p = .02$, $d = 0.12$, and the shifted tonal variation ($M = 0.28$, $SD = 0.17$), $t(23) = 2.07$, $p = .03$, $d = 0.12$, but not with the nonshifted tonal variation ($M = 0.24$, $SD = 0.17$), $t(23) = -1.68$, $p = .95$, $d = -0.07$. Results for CVs thus closely matched what was found for error rates.

Discussion

The results from Experiment 2 extended the finding of atonal and tonal asymmetric similarity (Bartlett & Dowling, 1988; Dowling & Bartlett, 1981) to a production context. Though this phenomenon has been demonstrated in a number of other domains (e.g., Karylowski & Skarzynska, 1992; Schimmack & Reisenzein, 1997), this is the first finding of asymmetric *perception–action similarity*. This is consistent with Hypothesis 2 in that serially shifting a melody of a different tonal class is only disruptive when planning an atonal melody due to a lack of pitch expectations in such melodies. Here, participants made more errors in conditions when feedback differed in tonal class (i.e., the main effect of tonal class), in contrast to Experiment 1 and Pfordresher (2008).

The implication of these results is that participants treated the contour-matched tonal sequences as similar to the planned atonal melody. Thus, a planned atonal melody may allow activation of hierarchical organizations that include tonal as well as atonal structures. By contrast, planning a tonal melody may not allow generalization to atonal structures. These results suggest that tonal class determines perception–action similarity but only unidirectionally: that is, only when a perceived atonal melody is compared to a planned tonal melody.

The most tenable explanation for this asymmetry of results is, as was mentioned before, that atonal sequences do not activate the kinds of strong melodic expectations that are characteristic of tonal music. Lacking these expectations allows a greater range of melodies (i.e., those that are tonal and matched in contour) to be perceived as similar.

In terms of significant effects, the recorded difficulty ratings were generally consistent with error rates. That is, both serially shifted feedback conditions were rated as more difficult than normal feedback and also led to an increase in errors. Furthermore, the effect sizes of these conditions on ratings were large (Cohen, 1988) and comparable to those of Experiment 1. However, one

---

[2] An anonymous reviewer questioned whether the fact that tonal feedback disrupted planned atonal melodies, but not the reverse, was an artifact of general task difficulty. That is, performing atonal sequences with any type of altered feedback may be simply more difficult than performing tonal melodies with altered feedback. To test this, we pooled the error rate data across experiments and compared the differences between the control conditions (viz., the planned tonal and atonal melodies) and their altered feedback conditions. The results showed that these altered feedback conditions were not generally more difficult when planning an atonal melody.

discrepancy was that the nonshifted tonal condition increased errors yet was not rated as being more difficult than the control condition. This was precisely opposite to what was found for comparable conditions in Experiment 1, where the nonshifted atonal feedback melody was rated as more difficult, yet did not increase errors. It is likely that both of these effects relate to the familiarity of tonal versus atonal structures. In Experiment 2, familiar-sounding tonal melodies do not seem difficult when matching the produced melody in contour, yet they may disrupt melodies because the same familiarity disrupts action planning.

Experiments 3 and 4

Experiments 1 and 2 further investigated the role of tonal class in perception–action associations during music performance. One limitation in both experiments was that the pitch range for the atonal feedback melodies was two semitones smaller than for tonal feedback (C#–F# vs. C#–G#). As a result, pitch range may have influenced the perceived similarity of feedback to planned melodies in addition to tonal class. Experiments 3 and 4 were designed to address this limitation by having designs based on Experiments 1 and 2 (respectively) but with the range of pitches that matched across tonal and atonal feedback melodies.

Method

Participants

Experiment 3. Twenty-four undergraduate students (11 men, 13 women) participated in exchange for credit toward an introductory psychology course. Participants reported $M = 2.17$ years of piano training (range = 0–14) and $M = 2.54$ years of musical training not related to the piano (range = 0–11). Ten participants (42% of the sample) reported at least one year of piano training (M for this subset = 5.2, max = 14), and of these, three participants (13% of the entire sample) reported more than 5 years (specific values = 7, 12, 14). Three participants reported being left-handed, and no participant reported hearing deficits or any impairment in the use of their dominant hand. None reported having absolute pitch. The mean age across participants was 19 years (range = 18–21).

Experiment 4. Twenty-four undergraduate students (10 men, 14 women) participated in exchange for credit. Participants reported $M = 0.75$ years of piano training (range = 0–6) and $M = 1.67$ years of nonpiano training (range = 0–8). Seven participants (29% of the sample) reported at least one year of piano training (M for this subset = 2.6, max = 6) with one participant reporting more than 5 years of training. One participant reported being left-handed, and no participant reported hearing deficits or any impairment in the use of their dominant hand. Again, none reported having absolute pitch. The mean age across participants was 19.45 years (range = 18–33).

Conditions, Materials, Procedure, and Analysis

The procedure and conditions of these two experiments were identical to Experiments 1 and 2 and differed only in key-pitch mapping for the atonal feedback melodies. The pitch range for atonal melodies spanned seven semitones and comprised the pitches C4#, D4, D4#, F4#, G4#. The keyboard-pitch mapping for tonal feedback sequences remained unchanged from the first two experiments.

Data from three participants in Experiment 4 were omitted from the analyses because of error rates that exceeded .25 (range = .25–.35) in the control condition (normal feedback). This indicated that these subjects failed to learn the melody properly, as more than one out of four keypresses were incorrect (i.e., at least two errors in each repetition of the melody). Pooling the data from the previous three experiments, the mean error rate under normal feedback was 0.048 (range = 0–.14), indicating that these three participants were outliers. In all other respects, the statistical analyses were identical to the previous two experiments.

Results (Experiment 3)

A two-way ANOVA on performance error rates (Figure 2C) yielded results consistent with Experiment 1. There was a main effect of serial shift, $F(1, 23) = 7.12$, $p = .014$, $\eta_p^2 = .24$, $\eta_G^2 = .02$, no significant effect of tonal class, ($p = .26$, $\eta_p^2 = .05$, $\eta_G^2 = .002$), and no significant interaction effect ($p = .61$, $\eta_p^2 = .008$, $\eta_G^2 < .001$). Planned comparisons indicated that relative to normal (tonal) feedback ($M = 0.054$, $SD = 0.037$), only the serially shifted tonal condition ($M = 0.068$, $SD = 0.040$) resulted in a significant increase in error rates, $t(23) = 2.60$, $p = .008$, $d = 0.37$. Neither the atonal condition ($M = 0.053$, $SD = 0.040$) nor the serially shifted version ($M = 0.062$, $SD = 0.043$) produced significantly higher error rates ($p = .397$, $d = -0.04$, and $p = .083$, $d = 0.19$, respectively). These results replicated the primary finding of Experiment 1 and Pfordresher (2008) in that when tonal melodies were planned, feedback melodies differing in tonal class were treated as irrelevant to the planned melody (similar to random feedback).

The analysis of difficulty ratings (Figure 3C) also replicated Experiment 1. There was a main effect of serial shift, $F(1, 23) = 19.24$, $p = < .001$, $\eta_p^2 = .46$, $\eta_G^2 = .21$, and tonal class, $F(1, 23) = 10.65$, $p = .005$, $\eta_p^2 = .32$, $\eta_G^2 = .02$, but no significant interaction effect ($p = .06$, $\eta_p^2 = .15$, $\eta_G^2 = .01$). The results of the planned comparisons were identical to Experiment 1, as subjects reported all altered feedback conditions as more difficult than the control ($M = 2.20$, $SD = 0.78$). This included the serially shifted tonal condition ($M = 3.34$, $SD = 1.07$), $t(23) = 4.41$, $p < .001$, $d = 1.23$, the serially shifted atonal condition ($M = 3.40$, $SD = 1.01$), $t(23) = 4.81$, $p < .001$, $d = 1.33$, and the nonshifted atonal condition ($M = 2.66$, $SD = 0.82$), $t(23) = 3.56$, $p = .001$, $d = 0.57$.

We also analyzed the mean and variability (CV) of produced IOIs during each trial (see Table 1). In contrast to Experiment 1, Experiment 3 yielded significant effects of feedback on mean IOI. Nevertheless, these effects deviated from error rates in important ways. For mean IOI, the ANOVA yielded a significant main effect for shift, $F(1, 23) = 13.01$, $p = .002$, $\eta_p^2 = .36$, $\eta_G^2 < .01$, but no interaction ($p = .69$, $\eta_p^2 = .01$, $\eta_G^2 < .001$), similar to error rates. However, unlike the analysis for error rates, there was also a main effect of tonal class, $F(1, 23) = 5.88$, $p = .02$, $\eta_p^2 = .20$, $\eta_G^2 < .001$. Planned comparisons likewise showed some further inconsistency with error rates. Unlike error rates, both contrasts involving serially shifted conditions were significantly different from the control ($M = 417.92$, $SD = 79.81$; although the effects were small): the shifted tonal melody (M = 432.23, SD = 89.87), $t(23) = 2.59$, $p = .008$, $d = 0.17$, and the shifted atonal variation (M = 428.71, SD = 85.82), $t(23) = 2.78$, $p = .005$, $d = 0.13$. The contrast with the atonal nonshifted condition was nonsignificant (M = 412.17, SD = 81.52), $t(23) = -1.68$, $p = .95$, $d = -0.07$.

The ANOVA on timing variability (CV IOI) yielded results that were more consistent with error rates. There was a significant main effect of shift, $F(1, 23) = 6.86$, $p = .02$, $\eta^2_{\mathrm{p}} = .23$, $\eta^2_{\mathrm{G}} = .01$, but no main effect of tonal class ($p = .20$, $\eta^2_{\mathrm{p}} = .07$, $\eta^2_{\mathrm{G}} < .001$), and no interaction ($p = .88$, $\eta^2_{\mathrm{p}} = .001$, $\eta^2_{\mathrm{G}} < .001$). Planned comparisons yielded a significant contrast between the control (M = 0.24, SD = 0.14) and tonal shifted condition (M = 0.26, SD = 0.12), $t(23) = 2.07$, $p = .03$, $d = 0.22$. However, the contrast between the serial shift of the atonal variation (M = 0.25, SD = 0.12) and the control condition (tonal, nonshifted) fell short of significance (p = .06, d = 0.15). As in the other analysis, the comparison with the other nonshifted condition (M = 0.23, SD = 0.14) was nonsignificant ($p = .21$, $d = -0.04$).

Results (Experiment 4)

A two-way repeated-measures ANOVA on error rates (Figure 2D) resulted in a main effect of serial shift, $F(1, 20) = 6.44$, $p = .02$, $\eta^2_{\mathrm{p}} = .24$, $\eta^2_{\mathrm{G}} = .028$, no main effect of tonal class ($p = .60$, $\eta^2_{\mathrm{p}} = .01$, $\eta^2_{\mathrm{G}} < .001$), and no interaction ($p = .71$, $\eta^2_{\mathrm{p}} = .006$, $\eta^2_{\mathrm{G}} < .001$). Thus, compared with Experiment 2, this omnibus ANOVA diverged only in the absence of a main effect for tonal class, possibly reflecting the control of pitch range in Experiment 4. Planned comparisons yielded results comparable to Experiment 2 for the serially shifted conditions. The serially shifted normal (atonal) feedback condition (M = 0.101, SD = 0.11) produced higher error rates relative to normal feedback (M = 0.070, SD = 0.047), $t(20) = 1.80$, $p = .044$, $d = 0.36$, as did the serially shifted tonal condition (M = 0.095, SD = 0.092), $t(20) = 1.81$, $p = .042$, $d = 0.34$. This implies that both shifted sequences maintained perception–action similarity to the originally planned melody, in support of Hypothesis 2 (asymmetric similarity). Unlike Experiment 2, here the nonshifted tonal condition (M = 0.069, SD = 0.074) did not result in significantly different error rates ($p = .45$, $d = -0.02$).

An ANOVA on the subjective difficulty ratings (Figure 3D) yielded only a main effect of serial shift $F(1, 20) = 12.81$, $p = .002$, $\eta^2_{\mathrm{p}} = 39.$, $\eta^2_{\mathrm{G}} = .11$. The main effect of tonal class was not statistically significant ($p = .25$, $\eta^2_{\mathrm{p}} = .07$, $\eta^2_{\mathrm{G}} = .002$), nor was the omnibus interaction effect ($p = .12$, $\eta^2_{\mathrm{p}} = .12$, $\eta^2_{\mathrm{G}} = .002$). Planned comparisons revealed that both the shifted atonal (M = 3.80, SD = 1.50) and shifted tonal condition (M = 3.80, SD = 1.32) were rated as more difficult compared to normal feedback (M = 3.03, SD = 1.24), $t(20) = 3.02$, $p = .007$, $d = 0.56$, and $t(20) = 3.11$, $p = .006$, $d = 0.60$, respectively. There was no significant difference between the nonshifted tonal condition (M = 2.79, SD = 1.07) and normal feedback ($p = .12$, $d = -0.20$). Thus, difficulty ratings replicated Experiment 2.

Means and CVs of IOIs, averaged across trials and participants, are shown in Table 1. For mean IOI, the ANOVA yielded a significant main effect for shift, $F(1, 20) = 5.72$, $p = .03$, $\eta^2_{\mathrm{p}} = .22$, $\eta^2_{\mathrm{G}} = .009$, but no main effect of tonal class ($p = .39$, $\eta^2_{\mathrm{p}} = .04$, $\eta^2_{\mathrm{G}} < .001$), and no interaction ($p = .83$, $\eta^2_{\mathrm{p}} = .002$, $\eta^2_{\mathrm{G}} < .001$), similar to error rates. The planned contrast between the nonshifted atonal control (M = 466.84, SD = 107.33) and the shifted tonal variation (M = 484.19, SD = 117.94) was significant, $t(20) = 1.77$, $p = .05$, $d = 0.15$, but curiously, the contrast with the shifted version of the planned melody (M = 486.58, SD = 121.10) fell short of significance ($p = .06$, $d = 0.17$). Not surprisingly, the contrast with the tonal nonshifted condition (M = 462.34, SD = 102.19) was not significant ($p = .79$, $d = -0.04$).

The ANOVA on timing variability (CV IOI) likewise yielded a significant main effect for shift, $F(1, 20) = 6.56$, $p = .02$, $\eta^2_{\mathrm{p}} = .25$, $\eta^2_{\mathrm{G}} = .01$, but no main effect of tonal class ($p = .86$, $\eta^2_{\mathrm{p}} < .01$, $\eta^2_{\mathrm{G}} < .001$), and no interaction ($p = .54$, $\eta^2_{\mathrm{p}} = .02$, $\eta^2_{\mathrm{G}} < .001$). Planned comparisons yielded a significant contrast between the nonshifted atonal condition (the control; M = 0.25, SD = 0.11) and both shifted conditions, the atonal shifted melody (M = 0.29, SD = 0.13), $t(20) = 2.33$, $p = .02$, $d = 0.28$, and the tonal shifted variation (M = 0.28, SD = 0.12), $t(20) = 1.78$, $p = .05$, $d = 0.22$, but the contrast with the tonal nonshifted condition (M = 0.26, SD = 0.12) was not significant ($p = .39$, $d = 0.03$).

Comparisons Across Experiments

A succinct comparison of Experiments 1 and 3 with Experiments 2 and 4 can be found in Table 2. This table highlights the most critical results across experiments: differences in error rates brought about by AAF manipulations of tonal class and serial shift. As can be seen, there were highly consistent results across all four experiments with respect to how planning a tonal or atonal melody influences one's susceptibility to disruption. Whereas serial shifts of the planned melody were always disruptive, serial shifts of a contour-matched variation only disrupted performers when the planned melody was atonal and the variation was tonal (Hypothesis 2). Effects of pitch range (which varied for atonal but not tonal feedback melodies) were represented in effect size magnitudes but were too subtle to qualify this pattern.

Discussion

Experiments 3 and 4 generally replicated the results Experiments 1 and 2, respectively, and ruled out the alternative explanation based on differences in pitch range across tonal and atonal feedback melodies. If pitch range had been a factor, it follows that the feedback melodies with a more similar pitch range (those in Experiments 3 and 4) would have led to greater perception–action similarity relative to Experiments 1 and 2. An increase in perception–action similarity would have caused more production errors when paired with the serial shift manipulation. However,

Table 2
Effect Sizes of Planned Comparisons for Error Rates Across Experiments (Exp.)

	Planned tonal		Planned atonal	
	Narrow pitch (Exp. 1): *d*	Wide pitch (Exp. 3): *d*	Narrow pitch (Exp. 2): *d*	Wide pitch (Exp. 4): *d*
Tonal class	−.06	−.04	.29*	−.02
Serial shift	.39*	.37**	.43**	.36*
Both	.28	.19	.83**	.34*

* $p < .05$. ** $p < .01$.

this was clearly not the case in terms of both the pattern of significant effects and effect sizes.

In terms of error rates, Experiment 3 replicated the primary result that participants produced more errors only when the planned tonal melody was serially shifted. Again, as in Experiment 1 (as well as Pfordresher, 2008, Experiment 3), the implication is that when planning a tonal melody, both the atonal and serially shifted atonal feedback were treated as dissimilar to the planned melody and did not interfere with action planning. Similarly, Experiment 4 replicated the primary results of Experiment 2, in which participants produced more performance errors (compared with normal feedback) when hearing either a serial shift of the planned atonal melody or a serially shifted version of the tonal variant. This implies that both melodies were treated as relevant to the planned melody, in support of Hypothesis 2. The subjective difficulty ratings in Experiments 3 and 4 mirrored their counterparts from Experiments 1 and 2.

Measures of produced tempo (mean IOI) did not comport with error rates entirely, reflecting to some degree the dissociation between sequencing and timing discussed earlier. In Experiment 3, both serial shift conditions slowed production, whereas this effect was specific to shifts of the planned melody for error rates. In Experiment 4, curiously, significant slowing was only found for serial shifts of the tonal variation. Thus, although the present results do not replicate the null effect of pitch alterations on timing reported elsewhere (e.g., Pfordresher, 2005), they nevertheless suggest different effects of these alterations on sequencing and timing in production. As noted earlier, timing variability can be more susceptible to disruption from serial shifts, with mixed results in past research (e.g., Pfordresher, 2003, reported disruption but with smaller effect size for CVs as opposed to error rates; Pfordresher, 2005 reported no effect of serial shifts on CVs). In Experiments 3 and 4, as in Experiment 2, timing variability matched what was found for error rates.

Finally, it can also be seen that the participants in Experiments 1 and 3 had a noticeably greater amount of piano training compared to Experiments 2 and 4 ($M = 1.75$ and $M = 2.17$ vs. $M = 0.88$ and $M = 0.75$ years of piano training, respectively). Using 5 years of training as a criterion, we identified seven musicians in Experiments 1 and 3 and just one across 2 and 4. This difference could have affected results, as musicians have shown a greater capacity to generalize across variations of a melody (Pfordresher, 2008). However, when analyses were performed with these individuals excluded, the results of the experiments did not change.[3]

General Discussion

The four experiments reported here demonstrated that performers are sensitive to relationships between the melody associated with a sequence of key presses (the planned melody) and the resulting set of pitches in auditory feedback (the feedback melody). Specifically, participants were influenced by the way in which the feedback melody's contour aligned with that of the planned melody (i.e., in serial shifts of feedback), but the tonal class of the planned and feedback melodies qualified this effect. If the planned melody was tonal and the feedback melody was atonal, serial shifts did not disrupt performance. However, when the planned melody was atonal, disruption resulted from a serial shift of a tonal feedback melody.

These results are significant in several respects. First, they demonstrate that tonal class is important for establishing the structural similarity of the feedback melody relative to the planned melody but that the similarity relationship between tonal and atonal classes is asymmetric (cf. Dowling & Bartlett, 1981; Tversky, 1977). Tonal melodies are restrictive. When planning a tonal melody, the performer generates schema-based expectations that restrict anticipated feedback events to those that construct a tonal melody, whereas expectations for atonal sequences may be indeterminate (Krumhansl, Sandell, & Sergeant, 1987). As a result, an atonal feedback melody is treated as independent of the tonal planned melody, and the performer is relatively immune from disruption when that feedback melody's contour is shifted. By contrast, atonal melodies are nonrestrictive, because they do not invoke an underlying memory schema that constraints expectations.

Second, these results are significant in that they appeared in a musically untrained population. A large number of past studies have shown that untrained listeners have implicit knowledge about tonality that influences perceptual expectations (Bigand & Poulin-Charronnat, 2006; Tillmann, Bharucha, & Bigand, 2000). It is more remarkable, however, that untrained individuals apply these tonal expectations to a keyboard production task. Consider the role of melodic contour. It is widely known that pitches on a keyboard are higher to the right and lower to the left. More broadly, sensitivity to melodic contour may reflect sensitivity to the trajectory formed by patterns in any modality (McDermott, Lehr, & Oxenham, 2008). Thus, it can be expected that nonpianists would be sensitive to contour relationships, which amount to mapping spatial relationships on the keyboard to expected pitch height. However, mapping of pitches to tonal versus atonal pitch sets involves a more refined knowledge of the keyboard that might be contingent on training, yet our nonpianists were sensitive to these relationships.

Third, these experiments shed further light on the relationship between performance and subjective awareness. Across experiments, mean error rate by condition was significantly correlated with mean difficulty ratings, $r(14) = .59$, $p < .01$. Yet important differences between the measures were found. A particularly informative result was found when examining the feedback conditions when tonal class was altered without the serial shift. When participants planned a tonal melody and heard an atonal feedback melody (Experiments 1 and 3), error rates did not increase, yet ratings of difficulty did. By contrast, difficulty ratings did not increase when participants planned an atonal melody and heard a tonal feedback melody (Experiments 2 and 4), whereas in Experiment 2 this manipulation did increase error rates. As a result, the correlation between difficulty ratings and error rates was zero for those conditions that did not include a serial shift of feedback ($N = 8$). In these cases, experienced difficulty may reflect unfamiliarity

---

[3] Across our samples, the number of musically trained participants was very small ($N = 8$) and unevenly distributed across experiments. However, when the musicians from Experiments 1 and 3 were analyzed separately, their results diverged from the nontrained participants in those studies: Instead of producing more errors only when the planned tonal melody was shifted, they also produced more errors in the shifted atonal condition ($d = 1.35$, $p = .01$). This provides some preliminary evidence that asymmetric perception–action similarity may be qualified by musical expertise.

rather than disruption. Somewhat paradoxically, the same unfamiliarity that increased difficulty ratings when hearing atonal AAF removed performers from the disruptive effect of serial shifts. Thus, as in previous research (Couchman, Beasley, & Pfordresher, 2012; Sato & Yasuda, 2005), the present data suggest that the subjective experience of AAF may be generated independently of the effect AAF has on motor planning.

Ultimately, these results add to a growing literature that suggests performers are sensitive to relationships between planned action sequences and the contents of auditory feedback. However, these relationships are based more on higher order sequential organization than specific item-level information. On the basis of this study and its predecessors (e.g., Pfordresher, 2005, 2008), the two aspects of structure that appear to be critical are melodic contour and tonal class. Certainly, other features are worth considering in future research, though it is likely that these two aspects of structure will remain dominant. Contour and tonal class are particularly salient in the recognition of newly learned melodies (Dowling, 1978), and contour has invariably been a significant factor of similarity in the perception domain (Dowling & Fujitani, 1971; Dowling & Harwood, 1986; McAdams, Vieillard, Houix, & Reynolds, 2004).

References

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, 37,* 379–384. http://dx.doi.org/10.3758/BF03192707

Bangert, M., & Altenmüller, E. O. (2003). Mapping perception to action in piano practice: A longitudinal DC-EEG study. *BMC Neuroscience, 4,* 26.

Bartlett, J. C., & Dowling, J. W. (1988). Scale structure and similarity of melodies. *Music Perception, 5,* 285–314. http://dx.doi.org/10.2307/40285401

Bigand, E., & Poulin-Charronnat, B. (2006). Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition, 100,* 100–130. http://dx.doi.org/10.1016/j.cognition.2005.11.007

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

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. http://dx.doi.org/10.1016/j.concog.2011.10.007

Cycling’74.Max6[ComputerSoftware]

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review, 85,* 341–354. http://dx.doi.org/10.1037/0033-295X.85.4.341

Dowling, W. J., & Bartlett, J. C. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology: A Journal of Research in Music Cognition, 1,* 30–49. http://dx.doi.org/10.1037/h0094275

Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval, and pitch recognition in memory for melodies. *The Journal of the Acoustical Society of America, 49,* 524–531. http://dx.doi.org/10.1121/1.1912382

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

Finney, S. A. (1997). Auditory feedback and musical keyboard performance. *Music Perception, 15,* 153–174. http://dx.doi.org/10.2307/40285747

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

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General, 141,* 2–18. http://dx.doi.org/10.1037/a0024338

Howell, P. (2004). Effects of delayed auditory feedback and frequency-shifted feedback on speech control and some potentials for future development of prosthetic aids for stammering. *Stammering Research, 1,* 31–46.

Karylowski, J. J., & Skarzynska, K. (1992). Asymmetric self-other similarity judgments depend on priming of self-knowledge. *Social Cognition, 10,* 235–254. http://dx.doi.org/10.1521/soco.1992.10.2.235

Keppel, G. (1991). *Design and analysis: A researcher's handbook* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Klockars, A., & Sax, G. (1986). *Multiple comparisons*. Thousand Oaks, CA: Sage.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. Oxford, England: Oxford University Press.

Krumhansl, C. L., & Cuddy, L. L. (2010). A theory of tonal hierarchies in music. In M. R. Jones, R. R. Fray, & A. N. Popper (Eds.), *Music perception* (pp. 51–88). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4419-6114-3_3

Krumhansl, C. L., Sandell, G. J., & Sergeant, D. C. (1987). The perception of tone hierarchies and mirror forms in twelve-tone serial music. *Music Perception, 5,* 31–77. http://dx.doi.org/10.2307/40285385

Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. *The Journal of Neuroscience, 27,* 308–314. http://dx.doi.org/10.1523/JNEUROSCI.4822-06.2007

Large, E. W. (1993). Dynamic programming for the analysis of serial behaviors. *Behavior Research Methods, Instruments, & Computers, 25,* 238–241. http://dx.doi.org/10.3758/BF03204504

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

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics, Doklady, 6,* 707–710.

McAdams, S., Vieillard, S., Houix, O., & Reynolds, R. (2004). Perception of musical similarity among contemporary thematic materials in two instrumentations. *Music Perception, 22,* 207–237. http://dx.doi.org/10.1525/mp.2004.22.2.207

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2008). Is relative pitch specific to pitch? *Psychological Science, 19,* 1263–1271. http://dx.doi.org/10.1111/j.1467-9280.2008.02235.x

Novembre, G., & Keller, P. E. (2011). A grammar of action generates predictions in skilled musicians. *Consciousness and Cognition, 20,* 1232–1243. http://dx.doi.org/10.1016/j.concog.2011.03.009

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods, 8,* 434–447. http://dx.doi.org/10.1037/1082-989X.8.4.434

Palmer, C., & van de Sande, C. (1993). Units of knowledge in music performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 457–470. http://dx.doi.org/10.1037/0278-7393.19.2.457

Palmer, C., & van de Sande, C. (1995). Range of planning in music performance. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 947–962. http://dx.doi.org/10.1037/0096-1523.21.5.947

Parncutt, R., Sloboda, J. A., Clarke, E. F., Raekallio, M., & Desain, P. (1997). An ergonomic model of keyboard fingering for melodic fragments. *Music Perception, 14,* 341–382. http://dx.doi.org/10.2307/40285730

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. http://dx.doi.org/10.1037/0096-1523.29.5.949

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. http://dx.doi.org/10.1037/0096-1523.31.6.1331

Pfordresher, P. Q. (2006). Coordination of perception and action in music performance. *Advances in Cognitive Psychology, 2,* 183–198. http://dx.doi.org/10.2478/v10053-008-0054-8

Pfordresher, P. Q. (2008). Auditory feedback in music performance: The role of transition-based similarity. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 708–725. http://dx.doi.org/10.1037/0096-1523.34.3.708

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. http://dx.doi.org/10.1111/j.1749-6632.2011.06408.x

Pfordresher, P. Q., & Beasley, R. T. E. (2014). Making and monitoring errors based on altered auditory feedback. *Frontiers in Psychology, 5,* 914. http://dx.doi.org/10.3389/fpsyg.2014.00914

Pfordresher, P. Q., & Kulpa, J. D. (2011). The dynamics of disruption from altered auditory feedback: Further evidence for a dissociation of sequencing and timing. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 949–967. http://dx.doi.org/10.1037/a0021435

Sato, A., & Yasuda, A. (2005). Illusion of sense of self-agency: Discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition, 94,* 241–255. http://dx.doi.org/10.1016/j.cognition.2004.04.003

Schimmack, U., & Reisenzein, R. (1997). Cognitive processes involved in similarity judgments of emotions. *Journal of Personality and Social Psychology, 73,* 645–661. http://dx.doi.org/10.1037/0022-3514.73.4.645

Sloboda, J. A., Clarke, E. F., Parncutt, R., & Raekallio, M. (1998). Determinants of fingering choice in piano sight-reading. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 185–203. http://dx.doi.org/10.1037/0096-1523.24.1.185

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review, 107,* 885–913. http://dx.doi.org/10.1037/0033-295X.107.4.885

Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327–352. http://dx.doi.org/10.1037/0033-295X.84.4.327

Yates, A. J. (1963). Delayed auditory feedback. *Psychological Bulletin, 60,* 213–232. http://dx.doi.org/10.1037/h0044155

Received April 9, 2015
Revision received October 2, 2015
Accepted October 3, 2015 ■