Performing Musical Dynamics: How Crucial are Musical Imagery and Auditory Feedback for Expert and Novice Musicians?

Laura Bishop
Austrian Research Institute for Artificial Intelligence (OFAI), Vienna, Austria

Freya Bailes
University of Hull, Hull, United Kingdom

Roger T. Dean
MARCS Institute, Penrith, Australia

Musicians anticipate and monitor the expressive effects of their actions during performance. Previous research suggests that the ability to imagine desired outcomes can partially compensate when auditory feedback is absent, permitting continued performance even though information about whether these outcomes are realized is unavailable. Research also suggests that musical imagery ability improves with increasing musical expertise. This study tested the hypothesis that expert musicians' superior imagery abilities enable reduced reliance on auditory feedback, relative to novice musicians, during the performance of loudness changes (i.e., dynamics). Musicians reproduced the dynamic changes of sounded scales using a loudness slider as the availability of imagery and auditory feedback was manipulated. Contrary to expectations, only novices showed impairment in performing dynamics during imagery disruption and auditory feedback deprivation. Experts showed limited dependence on both sources of information, suggesting greater flexibility in how musical information is mentally represented, compared to novices, and an improved ability to adapt planning strategies.

Received: May 21, 2013, accepted December 22, 2013.

Key words: musical imagery, auditory feedback, planning, musical expertise, expression

During performance, musicians can gauge how successfully they are achieving their expressive intentions by comparing the music they aim to produce with the music that is being produced. Specifications for future actions can be amended when planned and produced music differ (Furuya & Soechting, 2010; Keller & Koch, 2008; Palmer & Drake, 1997; Pfordresher, 2003, 2008). Both the expected and perceived effects of musicians' actions thus contribute to the successful production of an expressive performance. The present study investigated how reliant musicians are on these internal and external sources of information when performing musical dynamics (i.e., changes in loudness). Dynamics are among the most commonly manipulated parameters during expressive performance and have been shown to shape listeners' perceptions of sounded music in a number of ways, including influencing their perceptions of affect (Juslin, 2000), emotional valence and arousal (Schubert, 2004), and musical structure (Sloboda, 1985).

Musicians differ in the strategies they use to produce an expressive performance. Since not all musicians reliably achieve their expressive intentions, it is likely that some strategies are more effective than others. Expert musicians achieve their expressive plans with high precision (Palmer, 1997), which is likely attributable, in part, to effective performance planning (Ericsson & Lehmann, 1996; Keller & Koch, 2008; Palmer & Drake, 1997; Palmer & Meyer, 2000). An aim of the present study was to investigate whether expert musicians' superior planning abilities can generalize from familiar to novel motor contexts. Planning in music performance involves the activation in working memory of information about which actions need to be executed and what the effects of those actions should be (Keller & Koch, 2008; Palmer & Pfordresher, 2003). Performers develop expectations about what the effects of their actions should be as a result of their previous performance and listening experience (Bharucha, Curtis, & Paroo, 2006; Keller & Koch, 2008; Pfordresher, 2008). Expectations about events that are likely to occur in the context of a particular musical style are termed "schematic" (i.e., expectations about "how this kind of piece goes"), while expectations about upcoming events in a specific familiar sequence are termed "veridical" (i.e., expectations about "how this piece goes"). Veridical expectations that are highly accessible to conscious awareness can be described as musical imagery (Bailes & Delbé, 2009).

Music Perception, volume 32, issue 1, pp. 51–66, issn 0730-7829, electronic issn 1533-8312. DOI: 10.1525/mp.2014.32.1.51

MUSICAL IMAGERY AND ANTICIPATING THE EXPRESSIVE EFFECTS OF ACTIONS

Musical imagery is defined here as the conscious experience of music that is not an immediate consequence of its production or perception. Research has shown that imagined music can be analogous in form to sounded music, with tonal and temporal relationships maintained. Pitch (Aleman, Nieuwenstein, Bocker, & de Haan, 2000; Janata & Paroo, 2006; Pecenka & Keller, 2009), duration (Halpern, 1988; Janata & Paroo, 2006), and timbre (Halpern, Zatorre, Bouffard, & Johnson, 2004; Pitt & Crowder, 1992) can be maintained when music is imagined. Studies on whether loudness is imagined originally yielded conflicting results (Intons-Peterson, 1980; Pitt & Crowder, 1992; Wu, Yu, Mai, Wei, & Luo, 2011); however, recent research using continuous response methods has provided evidence that *loudness change*, or musical dynamics, can be imagined by both musicians and nonmusicians (Bishop, Bailes, & Dean, 2013a).

Previous research has shown imagery to be one of the processes involved in music performance planning (Bishop, Bailes, & Dean, 2013b; Keller, 2012; Keller, Dalla Bella, & Koch, 2010). In an fMRI experiment reported by Keller (2012), people tapped in synchrony with tempo-changing pacing signals while completing a working memory task that differed across conditions in its difficulty. Increased working memory demands led to a decline in prediction abilities – indicated by reduced success in synchronization – and were associated with decreased activity in several brain regions, including those implicated in auditory imagery and attention. These findings suggest that imagery may be important for the temporal coordination of actions. Imagery may also be important for the planning of expressive parameters such as dynamics (Bishop et al., 2013b; Repp, 1999; Rosenberg & Trusheim, 1990; Wöllner & Williamon, 2007). Bishop et al. (2013b) tested the hypothesis that consciously accessible imagery for dynamics and articulation can be used during piano performance to aid in the realization of expressive intentions. Pianists attempted to replicate their own expressive performance of two pieces as the presence of auditory and motor feedback was manipulated. Dynamic and articulation markings were introduced into the visually displayed scores as they played, and the pianists verbally judged whether each marking matched their expressive intentions while continuing to play with their own interpretation. Under both normal and absent feedback conditions, pianists produced accurate verbal responses faster than they would have if they had needed to wait until auditory feedback was available for retrospective assessment, suggesting that their plans included consciously accessible images of the dynamics and articulation that they intended to play.

Given the overlaps between sounded and imagined music, musical imagery in the context of performance could comprise the internal, anticipatory counterpart to what is actually performed and perceived as feedback. It has been predicted that imagery may compensate when auditory feedback is missing or unreliable (Highben & Palmer, 2004; Repp, 1999; Wöllner & Williamon, 2007). In the study by Bishop et al. (2013b), pianists replicated their dynamics and articulation successfully in the absence of auditory feedback. Verbal response times also shortened, suggesting that imagery strengthened and compensated for the lack of auditory feedback. These findings were somewhat in contrast to observations made by Repp (1999), who found that when skilled pianists were asked to perform a Chopin Étude with and without auditory feedback, the expressive timing and dynamics profiles they produced in the absence of feedback were slightly but significantly attenuated compared to those they produced under normal feedback conditions. Further investigation is thus needed to clarify how reliant pianists are on auditory feedback during expressive performance.

Highben and Palmer (2004) found that musicians with stronger aural skills performed memorized melodies more accurately following learning under auditory and motor feedback deprivation conditions than musicians with weaker aural skills. It was concluded that stronger aural skills might enable more effective imagery that can "fill in" for missing feedback. The extent of musicians' reliance on imagery during the performance of expressive parameters such as dynamics is not clear, however. Perhaps other planning processes (e.g., unconscious anticipation; nonimagery varieties of conscious planning, such as verbal self-instructions) can compensate for disrupted imagery just as imagery compensates in the absence of feedback. In the present study, the extent to which the performance of dynamics depends on contributions from musical imagery and the monitoring of auditory feedback was investigated. Musicians used a loudness slider to reproduce sounded dynamic profiles as the availability of imagery and feedback was manipulated. Their patterns of slider movements were compared with sounded dynamic profiles to provide a measure of how accurately they had reproduced dynamics from memory.

AUDITORY FEEDBACK AND MONITORING THE EXPRESSIVE EFFECTS OF ACTIONS

Monitoring is the process of comparing expected and perceived effects of actions and correcting for deviations between them (Palmer & Drake, 1997). A number of studies have examined the effects of sensory feedback disruption on performance as a way of investigating both monitoring processes and the way parameters such as pitch, melodic contour, timing, and loudness are specified in performance plans. These studies have shown that pitch accuracy tends to decline when the pitch of performed notes is altered (Couchman, Beasley, & Pfordresher, 2012; Furuya & Soechting, 2010; Keller & Koch, 2008; Pfordresher, 2003, 2008) and tempo tends to slow when the onsets of performed notes are delayed (Couchman et al., 2012; Kulpa & Pfordresher, 2012; Pfordresher, 2003).

Altering the loudness of performed notes has been shown to affect the force of subsequent movements, indicating that musicians are sensitive to deviations between expected and perceived loudness (Furuya & Soechting, 2010; Kunde, Koch, & Hoffman, 2004). The extent to which findings from these studies can generalize to expressive performance and the execution of musical dynamics is unclear, however. In both studies, the acoustic intensity of individual tones during a performance was altered. Such a manipulation is more representative of errors in technique, such as those that might result from a clumsy finger, than of a failure to perform dynamics as intended. Dynamics in the Western classical tradition sometimes constitute abrupt contrasts in loudness from one note to the next, but they can also involve gradual changes across one or more phrases, or contrasts between entire sections. Monitoring dynamics, therefore, involves not only detecting individual notes that are anomalous in their loudness, but also ensuring that larger-scale changes are carried out as intended. Further investigation of how crucial auditory feedback is for the performance of these large-scale variations in loudness could have implications for our understanding of how dynamic parameters of music are specified in performance plans. Moreover, in naturalistic contexts, musicians often encounter auditory feedback that differs from their expectations in some way, whether due to a new or suboptimal acoustic environment or an unfamiliar instrument. Their dependence on auditory feedback may contribute to how successfully they are able to achieve their expressive intentions under such conditions. It may be more important for musicians to focus on anticipating the effects of their actions when information about whether those effects have been achieved is lacking or unreliable. In the present study, musical imagery was expected to compensate, to at least some degree, for the absence of auditory feedback. It was hypothesized that participants' intended dynamics would be performed in the absence of auditory feedback more successfully if imagery could be used than if the use of imagery was prevented with a concurrent working memory task.

DO EXPERT MUSICAL PLANNING ABILITIES GENERALIZE TO NEW MOTOR CONTEXTS?

The present study assessed how accurately musicians could reproduce sounded musical dynamics using a simple, though novel, interface. In naturalistic performance situations, performers must simultaneously manipulate numerous interacting parameters, including pitch, timing, dynamics, and articulation (Juslin, 2000; Palmer, 1997; Sloboda, 1985). The complexity of the task and the number of variables present pose challenges for the controlled investigation of performers' planning abilities. In the present study, in contrast, the loudness slider enabled investigation of musicians' reliance on imagery and auditory feedback during the performance of a single dimension. It was expected that imagery would be used to guide the construction of action plans and aid in feedback monitoring during the dynamic reproduction task, just as it does in naturalistic performance situations. The simplified context provided the opportunity to address another question as well: whether expert planning abilities, previously observed primarily in well-practiced instrumental contexts (Highben & Palmer, 2004; Palmer & Meyer, 2000; Pfordresher, 2008; Takahashi & Tsuzaki, 2008), would generalize to the novel motor context introduced by the slider interface.

The ability to generalize expectations to a range of possible auditory effects seems to develop with increasing expertise (Palmer & Meyer, 2000; Pfordresher, 2008). Pfordresher (2008) altered the auditory feedback received by pianists and nonpianists performing action sequences on a keyboard, so that the melodies participants heard were either the same as they were performing or a transposed variation with identical contour, and were either sounded at the correct time or were serially shifted. Both pianists and nonpianists showed impaired performance for serially shifted original melodies, indicating that the altered feedback deviated from their expectations. While neither group showed significant impairment for transposed variations sounded at the correct time, pianists were more likely than nonpianists to show impairment for serially shifted transposed variations. These findings suggest that expectations for dynamic information, or information such as melodic contour that unfolds over time, may strengthen with music training. Palmer and Meyer (2000) suggested that planning may become more conceptual and more readily dissociated from motor specifications with increasing musical expertise. They investigated novice child and skilled adult pianists' abilities to transfer learning across conceptual and motor dimensions. Pianists learned simple melodies using a prescribed fingering, then learned corresponding melodies that either maintained or differed in contour and interval size (conceptual dimension), and either maintained or differed in fingering (motor dimension). The least experienced novices showed transfer by way of facilitated learning for the second set of melodies only when melodic relations and fingering were maintained, while more experienced novices showed transfer across both conceptual and motor dimensions. Skilled adults showed the greatest amount of transfer across conceptual dimensions, which suggests that performance plans may become increasingly abstract and conceptual as expertise develops.

While these studies provide evidence that skilled musicians can generalize expectations to a range of possible effects on their own instrument (i.e., within a single motor context), how readily they generalize expectations to novel motor contexts remains unclear. Skilled musicians are thought to anticipate the auditory effects of their actions more readily than novices in familiar motor contexts and be less reliant on auditory feedback as a result. If skilled musicians also anticipate the auditory effects of their actions more readily than novices in novel motor contexts, then they should be less reliant on auditory feedback than novices in novel motor contexts as well. In previous studies showing skilled musicians to rely less on auditory feedback than novices, participants have only been tested on their own instrument, however (Bishop et al., 2013b; Highben & Palmer, 2004; Takahashi & Tsuzaki, 2008). Assessing musicians' reliance on feedback in other motor contexts is necessary to determine whether skilled musicians' superior planning abilities are task-specific. Skilled musicians may even be more reliant on auditory feedback in novel motor contexts than novices. Auditory-motor associations have been shown to strengthen with increasing expertise, and auditory feedback that violates expectations arising from these associations tends to disrupt motor performance (Keller et al., 2010; Keller & Koch, 2008; Pfordresher, 2003, 2008; Pfordresher & Mantell, 2012). Skilled musicians might be subject to greater disruption when expectations arising from their strengthened auditory-motor associations are violated by the novel task.

In the present study, a loudness slider adjustment task was used to assess musicians' abilities to imagine and replicate previously sounded dynamic profiles. Unlike a typical music performance task (e.g., Repp, 1999), neither novices nor experts were expected to be familiar with the motor demands involved. If experts were found to rely more on anticipating or imagining the effects of their slider movements than novices, as was hypothesized, then this would indicate that expert performer imagery abilities are generalizable across motor contexts.

PRESENT RESEARCH

The aim of this study was to investigate the extent to which the performance of dynamics depends on contributions from musical imagery and the monitoring of auditory feedback. Musicians' abilities to reproduce sounded dynamic profiles under normal, disrupted imagery, and auditory feedback deprivation conditions were assessed. Some research suggests that auditory feedback contributes to the fine-tuning of expressive performance (Repp, 1999). On the other hand, expert musicians anticipate the effects of their actions more precisely than novices and may be less dependent on auditory feedback as a result (Bishop et al., 2013b; Highben & Palmer, 2004; Takahashi & Tsuzaki, 2008). In the present study, it was predicted that experts' superior imagery abilities would allow for increased use of anticipation and decreased reliance on auditory feedback relative to novices during the reproduction of sounded dynamics.

Novice and expert musicians listened to musical scales played with various dynamic changes, and then following a brief retention interval, reproduced each dynamic profile by moving a slider up to indicate a crescendo or down to indicate a decrescendo. This continuous response task was used instead of a recognition task (e.g., Wu et al., 2011) so that imagined *loudness change* could be assessed. Musical scales sounded with dynamic changes were used as stimuli in place of more naturalistic passages of music in order to discourage participants from relying on abstract knowledge of Western expressive conventions to complete the task and to instead encourage the use of imagery. The experiment comprised four conditions, completed in blocks by all participants: 1) a *control* condition in which imagery was undisturbed and sound was available as dynamics were reproduced, 2) an *auditory feedback deprivation* condition in which imagery was undisturbed but dynamics were reproduced in silence, 3) a *disrupted imagery* condition in which sound was available as dynamics were reproduced, but participants completed a simultaneous visual gesture interference task to disrupt imagery for dynamics, and 4) an *auditory feedback deprivation/disrupted imagery* condition in which participants completed the simultaneous visual gesture interference task and reproduced dynamics in silence.

The visual gesture interference task was developed in an earlier study (Bailes, Bishop, Stevens, & Dean, 2012). This task uses a modified dual-task paradigm in which two overlapping working memory tasks interfere with each other: one task requires the retention of musical dynamics, and the other requires the retention of visually presented conductor gestures. Working memory, which mediates mental imagery, can be conceptualized as a process by which chunks of information become more or less active in mind (Baddeley, 1992; Keller, 2012). Only small amounts of similar material can be kept active at a time. In concurrent working memory tasks requiring people to retain two sets of similar material simultaneously, memory interference can occur (e.g., Baddeley, 1992; Bailes et al., 2012; Keller, 2012). In the visual gesture interference task used in the present study, conductor gestures are viewed following the presentation of the dynamic scale. Gestures are expected to interfere with the rehearsal of dynamics, and imagined dynamics are expected to interfere with the encoding of visual gestures. Participants then use the slider to reproduce the dynamic profile. The mental rehearsal of visual gestures is expected to interfere with their use of online imagery, disrupting participants' abilities to construct action plans and fill in for missing auditory feedback. Action planning is likewise expected to interfere with the rehearsal of visual gestures. An earlier study (Bailes et al., 2012) tested the efficacy of this visual gesture interference task. Musicians' success at mapping out dynamic profiles under visual gesture, verbal, and tonal interference conditions was compared to their success at mapping out dynamic profiles without imagery interference. Only visual gesture interference was found to significantly impair musicians' abilities to recall dynamics, in line with research suggesting motor involvement in musical imagery (Halpern et al., 2004; Hickok, Buchsbaum, Humphries, & Muftuler, 2003; Mikumo, 1994) and hypotheses of an association between movement and musical dynamics (Dean & Bailes, 2010; Eitan & Granot, 2006; Kunde et al., 2004).

In the present study, it was hypothesized that both expertise groups would replicate dynamic profiles most reliably during the control condition, when imagery was undisturbed and auditory feedback accompanied slider movements. It was also hypothesized that both groups would reproduce dynamic profiles least reliably during the auditory feedback deprivation/disrupted imagery condition, when the visual gesture interference task impeded on participants' imagery for dynamic stimuli and slider movements were made in silence. Expert musicians were hypothesized to be less dependent on auditory feedback than novices and to have better musical imagery abilities, enabling better retention and reproduction of dynamic profiles in conditions without imagery disruption. It was not expected that experts would be able to compensate for imagery disruption as well as they could for auditory feedback deprivation. As a result, the difference between novice and expert groups in terms of the amount of impairment displayed was hypothesized to be greater in the auditory feedback deprivation condition than in the disrupted imagery condition, and negligible in the auditory feedback deprivation/disrupted imagery condition.

Method

PARTICIPANTS

Forty-nine musicians (32 female) completed the experiment. Twenty-four were "expert performers" who had at least five years of formal lessons and either played professionally or had a university degree in music, and 25 were "novice performers" with between one and five years of formal training, who did not play professionally or have a university degree in music. Experts reported a mean age of 30.7 years ($SD = 10.1$) and a mean 11.9 years ($SD = 6.3$) of music training; novices reported a mean age of 22.2 years ($SD = 9.0$) and a mean 3.1 years ($SD = 2.0$) of music training. The difference in age between the expert and novice groups was significant, $t(41) = 2.96$, $p < .01$, but is unlikely to have affected the results as age did not correlate significantly with performance on the dynamic reproduction task under any of the conditions (all $p > .05$).

All participants completed the Ollen Musical Sophistication Index (OMSI) (Ollen, 2006, 2009), which categorizes people as more or less musically sophisticated based on criteria such as the quantity and level of formal training, practice habits, composition experience, and concert attendance. Scores below 500 on the OMSI indicate people who are "less musically sophisticated" and scores above 500 indicate people who are "more musically sophisticated." Experts in the current study achieved a mean score of 701 ($SD = 196$), and novices achieved a mean score of 238 ($SD = 173$); this difference in OMSI score was significant, $t(40) = 8.4$, $p < .001$. Musicians played a variety of instruments and were recruited from universities, music societies, and music schools in the Sydney area. Undergraduate psychology students at the University of Western Sydney received course credit; all others received a small travel reimbursement.

All data for two experts were lost due to technical problems, and data for the first two participants tested (one expert and one novice) were excluded because alterations were subsequently made to the method of data collection.

STIMULI

Dynamic scales. Eighteen different dynamic profiles were created and superimposed on ascending/descending one-octave (16 notes) major scales. Six of these profiles were reserved for practice trials and training with the slider; the remaining 12 were superimposed on two scales each. In the Western classical music tradition, it is common for increases in pitch and increases in acoustic intensity to coincide (Eitan, 1997), and participants could have found remembering dynamic scale stimuli that violated this convention more difficult. The relationship between direction of pitch and loudness change was therefore controlled, and each dynamic profile was superimposed on one scale that began with ascending pitch and one that began with descending pitch. Dynamic profiles contained between two and four changes (crescendo or decrescendo), each ranging between three and eight notes in length. Dynamics were constrained so that two of the same type of change never occurred consecutively (e.g., one crescendo never immediately followed another). Three scales were created in each of eight different keys to make up the 24 dynamic stimuli used on experimental trials.

Scales were generated in real-time in a custom-made patch in Max/MSP, from pre-written specifications. Each note was 500 ms in length, so every scale spanned a total of 8 s. Pilot testing had showed that participants were unable to keep up in their slider movements when scales were presented at a faster pace, and unable to retain any of the dynamic or interference information for the duration of a trial when scales were presented at a slower pace. The same note velocity range was used for all dynamic changes (MIDI velocity range 20 – 60), and the same note velocity was never repeated by more than two consecutive notes. A single velocity was assigned to each sounded note. The virtual piano software PianoTeq was used in conjunction with Max/MSP to produce a naturalistic piano tone.

The scores viewed by participants on experimental trials were written in Sibelius 4 and contained key and time signatures and notes, but no dynamic markings. All were written in the treble clef. Scores were presented in Max/MSP and gradually revealed at the same pace as they had been sounded, so that notes appeared one after another every 500 ms. Examples of scores viewed by participants are shown in Figure 1.

Visual gesture sequences. Ten clips depicting conductor gestures were selected from the longer videos available in *Expressive Conducting* (Wiens, 2002). All clips featured a single conductor against a black background and were recorded from the rear right of the conductor, so that the face and left hand were not visible, but the white baton could be seen clearly at all times. The clips selected spanned a range of movement magnitudes and speeds, but no attempt was made to link conductor movements with scale dynamics or control for the potential relation between them. Each clip was 2 s in length; original clips that were shorter than 2 s (minimum 1.8 s) were stretched in Adobe Premiere Pro CS4.

Clips were combined randomly into 14 sequences. Twelve sequences were used on experimental trials, and two were reserved for practice trials. All gesture sequences were assembled in iMovie. Each sequence comprised three silent clips separated by 1-s views of a blank (black) screen. Three clips were used because research on working memory span for body movements suggests that people can remember an average of three movements at a time (Wachowicz, Stevens, & Byron, 2011; Wood, 2007). Clips were not repeated within individual sequences. For half the sequences, a corresponding sequence was constructed for test presentation in which one of the three clips was replaced by a different clip. Starting, middle, and end clips were replaced an equal number of times across these "different" sequences.

EQUIPMENT

The experiment was run from a MacBook (OS X 10.5.8), with stimuli presented and data collected in Max/MSP. An I-CubeX push v1.1 (100 mm in length) was used to collect slider data (Figure 1). The slider was fixed to a plastic box that inclined away from participants, so that movement upwards and away from them corresponded to an increase in loudness, and movement downwards and towards them corresponded to a decrease in loudness. Participants wore Koss UR-20 headphones throughout the experiment and made "same" and "different" responses using a labelled keypad.

DESIGN

A 2 (normal imagery; disrupted imagery) x 2 (normal auditory feedback; auditory feedback deprivation) within-subjects design was used, yielding four conditions that were completed as blocks. The order of conditions was pseudorandomized across participants, such that each condition was completed first by an equal number of people in each expertise group. Dynamic stimulus presentation was randomized across the entire experiment and each stimulus was presented only once. Visual gesture sequence stimulus presentation was

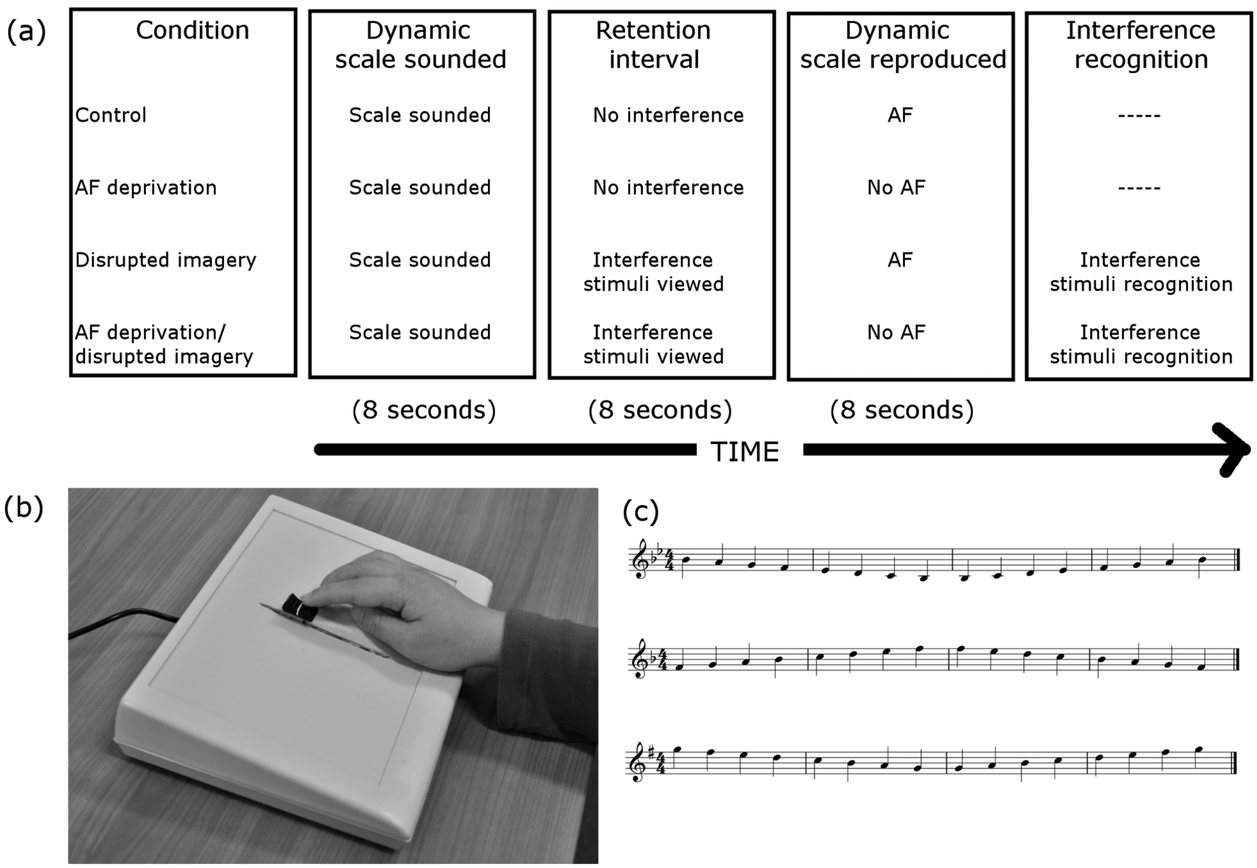


FIGURE 1. Experimental stimuli and procedure. (a) Structure of experimental trials under the four different conditions (AF = auditory feedback). (b) Photograph of the I-CubeX push v1.1 slider. Upwards slider movements indicated crescendi, and downwards slider movements indicated decrescendi. (c) Scores for three sample dynamic scales. Scales in eight different keys were constructed. Scores were revealed gradually, with one note appearing every 500 ms.

similarly randomized across the disrupted imagery trials and each was presented only once.

PROCEDURE

Participants began by filling out the OMSI questionnaire. They then completed a set of training trials to practice using the slider. Musical scores were presented in their entirety during the slider training phase and included dynamic markings. As the notes of the scale sounded, participants were to produce the marked dynamics by moving the slider up in response to crescendo markings and down in response to decrescendo markings. Notes sounded at a constant velocity unless the slider was moved. The experimenter observed the training trials and provided verbal feedback and further instructions as needed. Participants completed at least seven training trials. Upon reaching the end of the training, they were allowed to do the seven trials a second time if they did not yet feel confident with the task.

Following the slider training, participants received written and verbal instructions about the general format of the experiment. Additional instructions were provided at the start of each new block, followed by a practice trial. Each block consisted of six experimental trials and a practice. The entire experiment lasted about 45 min.

Figure 1 outlines the structure of the experiment trials completed during each condition. At the start of each trial, a dynamic scale was sounded, followed by an 8-s retention interval. In the *control* and *auditory feedback deprivation* conditions, participants heard nothing and saw a blank computer screen during this interval; during the *disrupted imagery* and *auditory feedback deprivation/ disrupted imagery* conditions, they viewed a sequence of conductor gesture clips during the interval. Participants were not required to reproduce the gestures overtly. Following the retention interval, a musical score corresponding to the previously sounded scale was revealed a note at a time on the screen, and participants used the slider to reproduce the dynamic profile for that scale. During the *control* and *disrupted imagery* conditions, the notes of the scale sounded a second time and participants' slider adjustments controlled the intensity of notes produced. As during the slider training, notes sounded at a constant velocity unless the slider was moved. During the *auditory feedback deprivation* and *auditory feedback deprivation/disrupted imagery* conditions, participants heard nothing and were asked to "replay" the scale in their heads while using the slider to indicate imagined dynamics. A 2-s orientation period preceded the visual display of scores in all conditions; this was indicated by a yellow circle that appeared on the screen with the words "Adjust slider." Participants were told to use this time to move the slider to a position representative of the starting dynamic of the scale. Following performance of the dynamic profile, a second sequence of conductor gestures was presented in *disrupted imagery* and *auditory feedback deprivation/disrupted imagery* conditions, which was either the same as the first sequence or differed by one clip. Participants pressed buttons on a computer keypad to indicate their "same" or "different" judgements. They were told to respond as quickly as possible and response times were recorded.

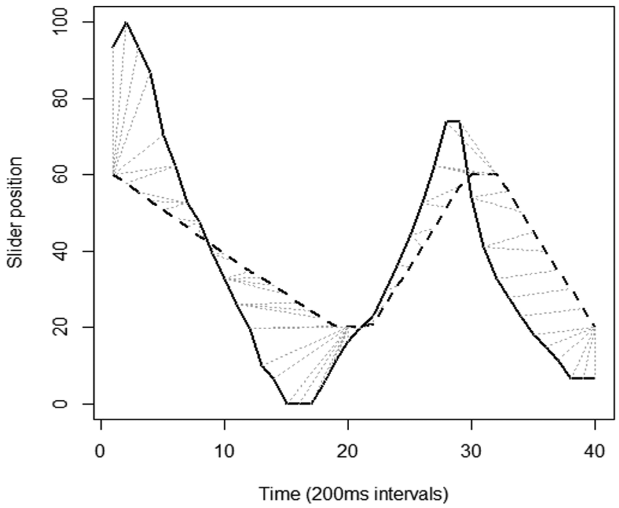
Slider data were recorded with a 200 ms sampling period, meaning that five slider positions were recorded for every two notes in the 16-note scales. This yielded a finely-sampled data series comprised of observations equally spaced in time. Other studies employing continuous response methods have used sampling periods of 250 ms (Luck, Toiviainen, & Thompson, 2010), 500 ms (Dean & Bailes, 2012; Geringer, 1995; McAdams, Vines, Vieillard, Smith, & Reynolds, 2004), or 1000 ms (Schubert, 2004); however, in the present study, less frequent sampling would have produced series that were representative of single loudness ratings for each individual note. More frequent sampling (e.g., every 100 ms) had the potential to produce data series with too much noise, and would not have been any more informative than the 200 ms period selected, since participants were not likely to be capable of indicating their loudness judgements at such a rapid pace or fine-grained level.

ANALYSIS OF DYNAMIC REPRODUCTION TASK SCORES

To assess how veridically participants reproduced dynamic profiles under the different conditions, dynamic time warping (DTW) (Bishop et al., 2013a; Giorgino, 2009) was used to calculate the distance between each participant response profile and the corresponding stimulus profile. In DTW, points that are most likely to correspond are identified along test (i.e., participant response) and reference (i.e., scale note velocity) data series. An average distance between profiles per event is then calculated that is independent of profile length. DTW can provide an accurate measure of similarity between data series that are autocorrelated (i.e., in which successive data points are not independent), which makes it suitable for use with time series data. It can also capture temporal misalignment between data series, which was important to consider in this study in order to differentiate participants who imagined the correct sequence of dynamic changes with the correct timing from participants who imagined the correct sequence of dynamic changes with incorrect timing.

The DTW algorithm allows for specifications to be made about how points on test and reference data series are aligned. In the present analyses, endpoint constraints were imposed, which meant that the first and last points on participant data series were always matched to the first and last points on reference data series. A symmetric step-pattern was used, so that multiple test series points could be mapped onto each reference series point, and multiple reference series points could be mapped onto each test series point. As slider position was sampled every 200 ms during the dynamic reproduction task, participant profiles were all 40 events in length (since trials were 8 s long). Reference key velocity profiles comprised one key velocity for every scale note, and were therefore 16 events in length. The use of a symmetric step-pattern was thus equivalent to comparing participant profiles to reference profiles that had been stretched to 40 events without interpolation between the key velocities that were actually sounded. Since DTW is sensitive to differences in magnitude between test and reference series curves, all participant data series were rescaled onto a slider range of 0 to 100 prior to computing the DTW distance. This allowed us to control for differences in the range of slider positions used by each participant across conditions. Sample participant response and scale key velocity profiles are shown in Figure 2.

Generalized linear modelling (GLM) was used to assess the effects of imagery and feedback conditions on the veridicality of reproduced dynamic profiles. The advantage of this method over ANOVA is that it is possible to specify the distribution of the data. The distribution of DTW distances was found to be highly skewed, so a gamma distribution was specified. Data from the 21 experts and 22 novices were pooled prior to conducting the analysis, and two outliers were removed as they fell more than 2.5 standard deviations away from the sample mean for their condition. One



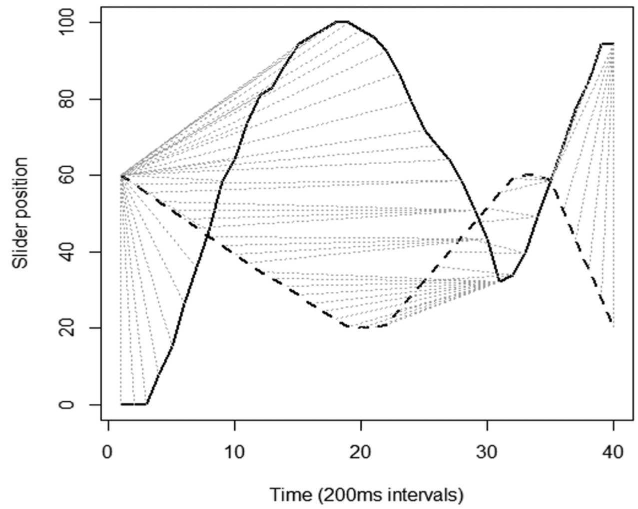


FIGURE 2. Sample participant response and scale key velocity profiles. Each plot shows the scale key velocity profile for one trial (dotted lines) and the response profile achieved by one participant (solid lines). The light grey lines indicate test and reference points that were matched by the DTW algorithm. In some instances, multiple test points were mapped to a single reference point or vice versa. Scale key velocity profiles have been stretched to span 40 events for the purpose of graphical display, and participant response profiles have been rescaled onto a range of 0 – 100, as they were for the calculation of DTW distances. Plot (a) depicts a dynamic profile recalled with high accuracy and temporal precision (DTW distance 5.8), while plot (b) depicts a dynamic profile recalled inaccurately (DTW distance 10.4).

outlier constituted an expert's data from the *auditory feedback deprivation/disrupted imagery condition*, and the other constituted another expert's data for the *disrupted imagery* condition. These participants' data for the other conditions were retained. The exclusion of individual cells does not preclude analysis of the rest of a participant's data in GLM, as one level of the specified independent variable is identified as the base, and all other levels are compared to that base. In the present analyses, the *control* condition was selected as the base, and dynamic reproduction task scores for each of the *auditory feedback deprivation*, *disrupted imagery*, and *auditory feedback deprivation/disrupted imagery* conditions were compared to scores for the *control*.

Results

EFFECTS OF IMAGERY DISRUPTION AND AUDITORY FEEDBACK DEPRIVATION

It was predicted that both imagery and auditory feedback would contribute to the successful reproduction of dynamics, with the least veridical performance being observed for the *auditory feedback deprivation/disrupted imagery* condition and the most veridical performance being observed for the *control* condition. To test this hypothesis, GLM was run on the distribution of DTW distances using condition and expertise group as predicting variables. The interaction between condition and group was not significant, $F(1, 166) = 0.91$, $p > .05$, but significant main effects of both factors were observed, $F(1, 167) = 4.42$, $p < .05$ (condition) and $F(1, 168) = 15.62$, $p < .001$ (group).

As predicted, participants reproduced dynamics most veridically in the *control* condition and least veridically in the *auditory feedback deprivation/disrupted imagery* condition (Figure 3). With data pooled across expertise groups, performance under each of the experimental conditions was compared to performance in the *control* condition to assess the effects of imagery disruption and auditory feedback deprivation. Differences between experimental conditions were not assessed, since the study was designed to investigate musicians' reliance on imagery and auditory feedack, not compare their reliance on one process relative to the other. Equating imagery and feedback manipulations was not within the scope of the study, so differences in performance under disrupted imagery and feedback deprivation conditions might have indicated differences in musicians' reliance on imagery and feedback, or they might have indicated differences in the strength of manipulation for each process. At a Bonferroni-adjusted alpha of .02, only performance during the *auditory feedback deprivation/disrupted imagery* condition was significantly worse than during the *control* condition, $t(41) = 3.51$, $p < .001$. Performance was worse during the *auditory feedback deprivation*, $t(42) = 1.64$, $p > .02$, and *disrupted*

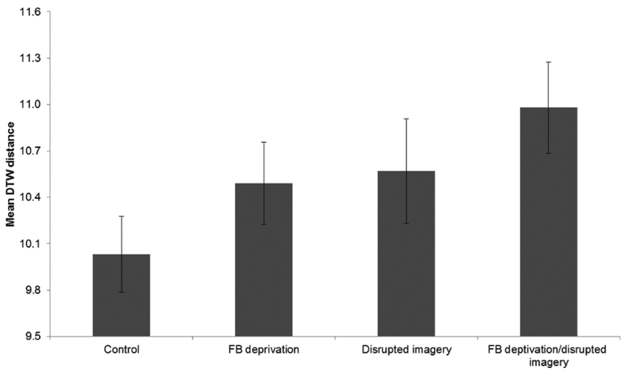


FIGURE 3. Dynamic reproduction task performance across conditions. Performance on the dynamic reproduction task was measured in terms of the distance between participant response profiles and the corresponding dynamic scale profiles, so a smaller DTW distance indicates better performance (FB = feedback). Error bars represent standard error of the mean.

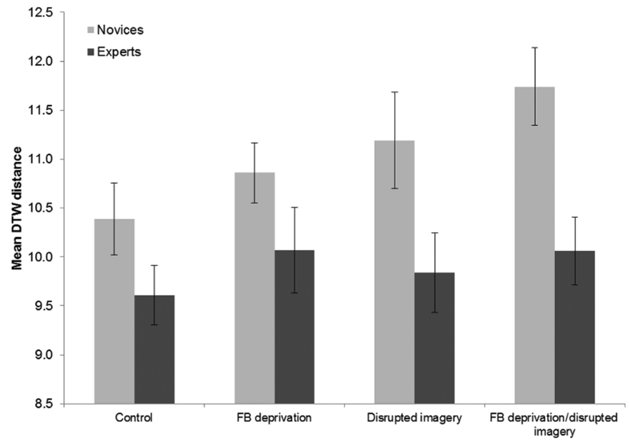


FIGURE 4. Dynamic reproduction task performance across conditions for experts and novices (FB = feedback). Error bars represent standard error.

imagery conditions, $t(41) = 1.42$, $p > .02$, than during the *control*, but neither difference was significant. This suggests that participants were better able to compensate for auditory feedback deprivation and imagery disruption when these sources of information were interrupted individually than when they were interrupted simultaneously.

EFFECTS OF MUSICAL EXPERTISE

It was hypothesized that the difference between expert and novice performance would be greater for the *auditory feedback deprivation* condition than the *disrupted imagery* condition, and only slight for the *auditory feedback deprivation/disrupted imagery* condition; experts were predicted to have superior imagery abilities that could compensate for a lack of auditory feedback as long as imagery was not also disrupted. The main effect of group, as stated above, indicates that experts outperformed novices in all conditions (Figure 4).

To assess whether the degree of difference between experts and novices differed between conditions, GLM was conducted on the DTW distances achieved under each condition separately, with expertise group as the independent variable. Experts significantly outperformed novices only during the *auditory feedback deprivation/disrupted imagery* condition, $F(1, 40) = 8.97$, $p < .01$, at a Bonferroni-adjusted alpha of .02. There was no effect of expertise group for the *control*, $F(1, 41) = 2.61$, $p > .02$, *auditory feedback deprivation*, $F(1, 41) = 2.05$, $p > .02$, or *disrupted imagery* conditions, $F(1, 40) = 3.26$, $p > .02$. Figure 4 suggests that experts performed similarly under all experimental conditions, while novices performed substantially worse when auditory feedback and imagery were made simultaneously unavailable. Thus, the significant difference between expert and novice performance observed during the *auditory feedback deprivation/disrupted imagery* condition may have resulted from novices' poor performance during that condition.

T-tests used to compare DTW distances within groups, across experimental conditions, provide evidence that only novices were substantially impaired by the combined absence of auditory feedback and disruption of imagery. Performance during *auditory feedback deprivation*, *disrupted imagery*, and *auditory feedback deprivation/disrupted imagery* conditions were compared to performance during the *control* condition for each expertise group individually. At a Bonferroni-adjusted alpha of .02, only the difference between novice *auditory feedback deprivation/disrupted imagery* performance and novice *control* performance was significant. Novices did not perform significantly worse during either *auditory feedback deprivation* or *disrupted imagery* conditions than during the *control* condition (all $p > .02$). Expert performance did not differ significantly from the *control* under any of the experimental conditions ($p > .02$). These findings are at odds with the hypothesis that experts would outperform novices most substantially when only auditory feedback was absent and experts' superior imagery abilities could enable enhanced compensation. Instead, the results suggest that expert musicians are not only non-reliant on auditory feedback in performing dynamics, but may be less reliant than novice musicians on imagery as well.

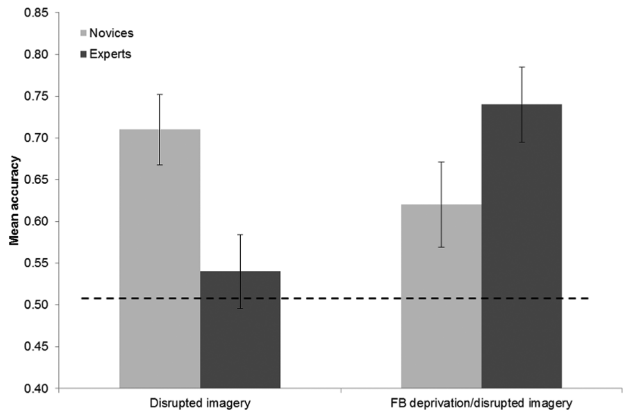


FIGURE 5. Accuracy of recognition judgements on visual gesture interference task for experts and novices. Error bars represent standard error and the dashed line represents chance performance (FB = feedback).

RECOGNITION OF INTERFERENCE STIMULI

The accuracy of recognition judgements made for interference stimuli used in both disrupted imagery conditions was calculated as the proportion of correct responses out of all responses given. Mean accuracy scores achieved by experts and novices during the *auditory feedback deprivation/disrupted imagery* and *disrupted imagery* conditions are shown in Figure 5. A two-way repeated measures ANOVA was run on the distribution of mean accuracy scores, using expertise group and condition as factors. There was no main effect of either group, $F(1, 80) = 0.50$, $p > .05$, or condition, $F(1, 80) = 0.06$, $p > .05$, but the interaction between them was significant, $F(1, 80) = 9.03$, $p < .01$.

Post-tests were conducted to investigate the relationship between expertise and recognition task performance in the two conditions. At a Bonferroni-adjusted alpha of .03, the difference between expert and novice performance during the *auditory feedback deprivation/disrupted imagery* condition was not found to be significant, $t(40) = 1.56$, $p > .03$; however, experts performed significantly worse than novices during the *disrupted imagery* condition, $t(41) = 2.89$, $p < .01$. Comparisons within-groups and across conditions showed that novices' recognition of interference stimuli did not differ between conditions, $t(21) = 1.31$, $p > .03$, but experts performed significantly more accurately during the *auditory feedback deprivation/disrupted imagery* condition than during the *disrupted imagery* condition, $t(20) = 2.91$, $p < .01$. Experts and novices may have differed in the strategies they used to complete the interference task. Though their recognition of interference stimuli was better during the *auditory feedback deprivation/disrupted imagery* condition, experts did not show a corresponding decline in performance on the dynamic reproduction task. In other words, their improved recognition of interference stimuli did not come at the cost of reproducing dynamics.

Discussion

Musicians imagine (Bishop et al., 2013b; Keller, 2012; Keller et al., 2010) and monitor (e.g., Palmer & Drake, 1997; Pfordresher, 2003, 2008) the effects of their actions during performance. The question of how crucial these internal and external sources of information are to the performance of expressive parameters such as dynamics, however, has received little attention in the literature. The present study assessed expert and novice musicians' dependence on imagery and auditory feedback with a task that required them to reproduce the dynamic profiles of sounded scales using a slider, under normal imagery and feedback, disrupted imagery, and auditory feedback deprivation conditions. The results were only partially in line with the hypotheses. As expected, the veridicality of reproduced dynamics declined most substantially when both auditory feedback and imagery were made simultaneously unavailable. However, contrary to expectations, only novices showed this pattern; experts showed no significant impairment under any of the conditions. Experts even recognized interference stimuli more accurately during the auditory feedback deprivation/disrupted imagery condition than during the disrupted imagery condition, despite showing similar performance on the dynamic reproduction task. Expert musicians' superior imagery abilities might have made them somewhat resistant to interference from the visual gestures. Experts across domains structure relevant information more effectively in working memory than novices, and can retain more material at a given time as a result (Chase & Simon, 1973; Williamson, Baddeley, & Hitch, 2010). Alternatively, expert musicians may not rely on either auditory feedback or imagery to perform dynamics. Rather, they may adapt their planning processes to task constraints more readily than novices and may be able to capitalize on alternate planning strategies. The current study was not designed to test what these alternate planning strategies might be, but below, we outline some possibilities that may merit further investigation.

ADAPTABILITY OF MUSICAL IMAGERY AMONG EXPERT PERFORMERS

Thought to contribute to expressive music performance is "musical intuition," a process in which long-term knowledge is reactivated without performers' conscious awareness or control, eliciting feelings that can be used in making decisions about interpretation and expression (Bangert, Schubert, & Fabian, 2009; Betsch, 2008). Some research suggests that musicians who rely on intuition to achieve their expressive intentions may achieve those intentions less successfully than musicians who are consciously aware of how they want the music to sound (Woody, 1999). Other research suggests that many of the decisions that skilled musicians make regarding interpretation and expression may be based on intuition (Bangert et al., 2009). In the present study, experts might have relied on an intuitive strategy for simultaneously retaining dynamics and conductor gestures. Novices showed no evidence of using a similar strategy to compensate for the simultaneous absence of feedback and disruption of imagery; rather, their performance during the auditory feedback deprivation/disrupted imagery condition was significantly worse than their performance during the control condition. The ability to retain and reproduce dynamics automatically may be a skill that develops with expertise.

An alternative explanation is that experts could have adopted the use of a different imagery technique to complete the task than was used by novices. Experts might have used descriptive imagery to retain dynamics, for instance. Descriptive imagery, or the mental representation of objects and events using a symbolic, language-like system, has been contrasted to depictive imagery, in which represented entities are similar in form to their physical analogues and spatial and temporal relationships are preserved (Pylyshyn, 1981; Thompson, Kosslyn, Hoffman, & Van der Kooij, 2008). In many musical contexts, it is likely that depictive and descriptive imagery are used in combination. Some situations may encourage reliance on depictive imagery while others encourage reliance on descriptive imagery. In the present study, for instance, the simultaneous absence of auditory feedback and disruption of imagery might have prompted experts to retain dynamics using a descriptive strategy. They might have encoded dynamics as a series of verbal descriptions that specified the direction and timing of loudness change (e.g., "decrescendo from notes one to three; crescendo from notes four to eight"), and then translated those descriptions into motor commands. As in other continuous response tasks, however, it is unlikely that a descriptive strategy would have been an effective way to retain and reproduce both the direction and the timing of dynamics accurately, given the pace at which scales were sounded and dynamics had to be reproduced (Bishop et al., 2013a, Lucas, Schubert, & Halpern, 2010).

A more effective strategy might have involved retaining dynamics and conductor gestures using contrasting types of depictive imagery. Expert musicians might have benefited from an enhanced ability to translate between different types of depictive imagery. For example, a musician who used a combination of auditory and motor imagery to retain dynamics during the control condition might have taken a primarily motoric approach to rehearsing conductor gestures and a primarily visual approach (e.g., using a mental image of an annotated musical score) to reproducing dynamics during the disrupted imagery conditions. Participants' music-reading abilities were not expected to influence their performance on the dynamic reproduction task, given the simplicity of the scores and the fact that only dynamics had to be reproduced. However, experts were likely more experienced than novices at reading music, and more experienced at maintaining visual images of musical scores in working memory (e.g., when playing from a score and needing to look at their instrument or glance up at a conductor or co-performer). Enhanced visual imagery abilities for musical material could have facilitated the simultaneous retention of visual gestures and dynamics.

The ability to adopt different strategies for imagining musical material would be suggestive of flexibility in musical imagery. Imagining dynamics may not require contributions from any particular modality, but rather, may depend on the context and the demands of the task (Intons-Peterson, 1980). The ability to translate material presented visually into a primarily auditory image, material presented acoustically into a primarily motor image, or material previously performed in silence into an auditory image is likely widespread among both musicians and nonmusicians, but flexibility in imagining music may improve with increasing expertise. If flexibility is a quality of expert musicians' imagery in particular, it might facilitate their expressive performance under conditions in which situational factors such as unexpected environmental distractors, the constraints of a new instrument, or the behavior of other musicians in an ensemble could act as sources of auditory, visual, or motor interference.

This possibility aligns with hypotheses made about internal models and their contributions to performance on perceptual-motor tasks. Two types of internal models – forward and inverse – are thought to represent the relationships between motor commands and their effects on the environment. Forward models enable prediction about the effects particular actions will have, while inverse models facilitate selection of the motor commands necessary to produce particular outcomes

(Keller, 2012). Some theoretical frameworks posit that multiple forward and inverse model pairings exist in the brain. With experience, these pairings become tightly coupled and specialized for certain contexts, though output from several model pairings can also be integrated to enable performance in novel situations (Haruno, Wolpert, & Kawato, 2001). In the current study, perhaps expert musicians had more highly developed networks of internal models than novices and were better able to switch between active models when the task demands changed.

DOES AUDITORY FEEDBACK DISTRACT OR FACILITATE PERFORMANCE IMAGERY?

Auditory feedback could be predicted to act as a distraction in some contexts, increasing cognitive load and making performance more demanding than it would be under auditory feedback deprivation conditions. Couchman et al. (2012) hypothesized that auditory feedback might interfere with performance plans when it deviates from performers' expectations in a way that makes the performers unclear about whether sounds can be attributed to their own actions. In the present study, the note intensities present in auditory feedback often differed in magnitude from those present in the original sounded dynamic scales, even when participants were reproducing dynamics correctly. The range of note intensities that could be produced by adjusting the slider (MIDI velocities 0 – 127) was substantially larger than the range of note intensities present in the dynamic scales (20 – 60). Participants were told to focus on reproducing the direction of dynamic changes rather than the magnitude of those changes, and many of them used nearly the full range of the slider in completing the task. For participants who attempted to reproduce dynamics during the control and disrupted imagery conditions using a different range of note intensity values than was present in the original dynamic scale stimuli, auditory feedback could have been sufficiently dissimilar to their expectations that it interfered with their plans. If this were the case, however, participants would have been expected to perform better in the absence of auditory feedback than in its presence, which was not the case. Thus, the presence of auditory feedback seems to have facilitated rather than impaired performance.

GENERALIZABILITY OF EXPERT PLANNING ABILITIES ACROSS MOTOR CONTEXTS

The results of the present study support previous findings of a relationship between musical expertise and the veridicality of imagined dynamics (Bishop et al., 2013a). The results also suggest that experts' superior planning abilities are generalizable to new motor contexts. Expert musicians may anticipate the effects of their actions more accurately and precisely than novices when performing on an instrument with which they are familiar, but in the present study, neither experts nor novices were expected to be familiar with the slider adjustment task. Both groups could be expected to associate upwards slider movements with increases in loudness. Furthermore, the task was expected to discourage musicians from encoding dynamics in terms of the familiar actions they would use when performing dynamics on their instrument or singing. The observed effects of expertise, then, suggest that experts' action planning abilities are not dependent on their technical skills. Musical imagery ability and the ability to predict the effects of well-practiced actions with precision are likely related but distinct capacities.

Similar observations have been made in other contexts in the literature. Musicians have been found to associate low pitch with a left response location and high pitch with a right response location more strongly than nonmusicians (Lidji, Kolinsky, Lochy, & Morais, 2007; Rusconi, Kwan, Giordano, Umiiltà, & Butterworth, 2005), for instance. Referred to as the spatial pitch association of response codes (SPARC) or spatial musical association of response codes (SMARC), this SPARC/SMARC effect is observed not only for pianists, who could be expected to map pitch height onto a horizontal spatial coding as a result of their keyboard experience, but non-pianist musicians as well. Keller and Koch (2008), likewise, observed a relationship between musical experience and the ability to imagine pitch contour in the context of a task that required participants to execute prescribed patterns of button-presses in response to visual cues. While enhanced performance might have been expected for pianists, effects of experience were observed among performers of a variety of instruments. The results of Keller and Koch's (2008) study suggest that associations between direction of movement trajectory and pitch contour strengthen with increasing musical expertise, regardless of the instrument played. The current results, similarly, suggest that associations between movement and dynamic change may strengthen with increasing musical expertise. In the current experiment, such strengthened associations might have facilitated the retention of dynamic scales and the construction of musical images.

Conclusions

A wide range of processes contribute to music performance planning; some are accessible to conscious awareness and control, while others are largely or entirely automatic. Among the aspects of planning accessible to consciousness is musical imagery, which skilled musicians report relying on in order to achieve their expressive intentions (Rosenberg & Trusheim, 1990). The present study investigated how crucial musical imagery and auditory feedback are for expert and novice musicians to reproduce sounded dynamics. The results suggest that novices are impaired by the disruption of imagery and absence of auditory feedback, while experts are dependent on neither. Previous research has shown that musical intuition, a form of planning that is not under conscious control, may contribute to expressive performance among skilled musicians (Bangert, et al., 2009). In the present study, experts may have adopted a more automatic strategy to performing dynamics when imagery and auditory feedback were unavailable. Experts may also be better able to adapt their planning strategies to novel task constraints, and they may be more flexible in how they imagine dynamic material. Future research should investigate the extent to which greater automaticity and flexibility in the mental representation of dynamic material can account for expert musicians' superior performance planning abilities. How these factors contribute to performance planning under more naturalistic circumstances, particularly in the context of suboptimal performance environments (e.g., in the presence of external noise) should also be investigated.

Author Note

This research was completed while Laura Bishop was a doctoral student and Freya Bailes was a Senior Research Fellow at the MARCS Institute, University of Western Sydney, Australia.

Correspondence concerning this article should be addressed to Laura Bishop, Austrian Research Institute for Artificial Intelligence, Freyung 6/6, A-1010 Vienna, Austria. E-mail: laura.bishop@ofai.at

References

Aleman, A., Nieuwenstein, M., Bocker, K., & de Haan, E. (2000). Music training and mental imagery ability. *Neuropsychologia, 38*, 1664-1668.

Baddeley, A. D. (1992). Working memory. *Science, 255*, 556-559.

Bailes, F., Bishop, L., Stevens, C. J., & Dean, R. T. (2012). Mental imagery for musical changes in loudness. *Frontiers in Perception Science, 3*. doi: 10.3389/fpsyg.2012.00525

Bailes, F., & Delbé, C. (2009). Long-term melodic expectation: The unexpected observation of distant priming effects. *Musicae Scientiae, 13*, 315-336.

Bangert, D., Schubert, E., & Fabian, D. (2009, December). *Decision-making in unpracticed and practiced performances of Baroque violin music*. Paper presented at the Second International Conference on Music Communication Science, Sydney, Australia.

Betsch, T. (2008). The nature of intuition and its neglect in research on judgment and decision making. In H. Plessner, C. Betsch, & T. Betsch (Eds.), *Intuition in judgment and decision making* (pp. 3-22). New York: Taylor and Francis.

Bharucha, J., Curtis, M., & Paroo, K. (2006). Varieties of musical experience. *Cognition, 100*, 131-172. doi: 10.1016/j.cognition.2005.11.008

Bishop, L., Bailes, F., & Dean, R. T. (2013). Musical expertise and the ability to imagine loudness. *PLOS ONE, 8*(2), e56052.

Bishop, L., Bailes, F., & Dean, R. T. (2013). Musical imagery and the planning of dynamics and articulation during performance. *Music Perception, 31,* 97-116.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4,* 55-61.

Couchman, J. J., Beasley, R., & Pfordresher, P. (2012). The experience of agency in sequence production with altered auditory feedback. *Consciousness and Cognition, 21*, 186-203. doi: 10.1016/j.concog.2011.10.007

Dean, R. T., & Bailes, F. (2010). A rise-fall temporal asymmetry of intensity in composed and improvised electroacoustic music. *Organised Sound, 15*, 147-158.

Eitan, Z. (1997). *Highpoints: A study of melodic peaks*. Philadelphia, PA: University of Pennsylvania.

Eitan, Z., & Granot, R. Y. (2006). How music moves: Musical parameters and listeners' images of motion. *Music Perception, 23*, 221-247.

Ericsson, K., & Lehmann, A. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology, 47*, 273-305.

Furuya, S., & Soechting, J. F. (2010). Role of auditory feedback in the control of successive keystrokes during piano playing. *Experimental Brain Research, 204*(2), 223-237. doi: 10.1007/s00221-010-2307-2

Geringer, J. M. (1995). Continuous loudness judgements of dynamics in recorded music excerpts. *Journal of Research in Music Education, 43*, 22-35.

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in R: The dtw package. *Journal of Statistical Software, 31*(7), 1-24.

Halpern, A. (1988). Mental scanning in auditory imagery for songs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 434-443.

Halpern, A., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia, 42*, 1281-1292. doi: 10.1016/j.neuropsychologia.2003.12.017

Haruno, M., Wolpert, D., & Kawato, M. (2001). MOSAIC model for sensorimotor learning and control. *Neural Computation, 13*, 2201-2220.

Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in Area Spt. *Journal of Cognitive Neuroscience, 15*, 673-682.

Highben, Z., & Palmer, C. (2004). Effects of auditory and motor mental practice in memorized piano performance. *Bulletin of the Council for Research in Music Education, 156*, 58-65.

Intons-Peterson, M. (1980). The role of loudness in auditory imagery. *Memory and Cognition, 8*, 385-393.

Janata, P., & Paroo, K. (2006). Acuity of auditory images in pitch and time. *Perception and Psychophysics, 68*, 829-844.

Juslin, P. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1797-1813.

Keller, P. E. (2012). Mental imagery in music performance: Underlying mechanisms and potential benefits. *Annals of the New York Academy of Sciences, 1252*, 206-213. doi: 10.1111/j.1749-6632.2011.06439.x

Keller, P. E., Dalla Bella, S., & Koch, I. (2010). Auditory imagery shapes movement timing and kinematics: Evidence from a musical task. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 508-513. doi: 10.1037/a0017604

Keller, P. E., & Koch, I. (2008). Action planning in sequential skills: Relations to music performance. *The Quarterly Journal of Experimental Psychology, 61*, 275-291. doi: 10.1080/17470210601160864

Kulpa, J. D., & Pfordresher, P. (2012). Effects of delayed auditory and visual feedback on sequence production. *Experimental Brain Research, 224*(2), 69-77. doi: 10.1007/s00221-012-3289-z

Kunde, W., Koch, I., & Hoffman, J. (2004). Anticipated action effects affect the selection, initiation, and execution of actions. *The Quarterly Journal of Experimental Psychology, 57A*, 87-106.

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

Lucas, B. J., Schubert, E., & Halpern, A. R. (2010). Perception of emotion in sounded and imagined music. *Music Perception, 27*, 399-412. doi: 10.1525/MP.2010.27.5.399

Luck, G., Toiviainen, P., & Thompson, M. R. (2010). Perception of expression in conductors' gestures: A continuous response study. *Music Perception, 28*, 47-57. doi: 10.1525/mp.2010.28.1.47

Mikumo, M. (1994). Motor encoding strategy for pitches and melodies. *Music Perception, 12*, 175-197.

Ollen, J. E. (2006). *A criterion-related validity test of selected indicators of musical sophistication using expert ratings* (Unpublished doctoral dissertation). The Ohio State University, Columbus, Ohio.

Ollen, J. E. (2009, August). *Cross-validation of a model for classifying musical sophistication*. Paper presented at the Society for Music Perception and Cognition Conference, Indianapolis, IN.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48*(1), 115-138.

Palmer, C., & Drake, C. (1997). Monitoring and planning capacities in the acquisition of music performance skills. *Canadian Journal of Experimental Psychology, 51*, 369-384.

Palmer, C., & Meyer, R. (2000). Conceptual and motor learning in music performance. *Psychological Science, 11*, 63-68.

Palmer, C., & Pfordresher, P. Q. (2003). Incremental planning in sequence production. *Psychological Review, 110*, 683-711. doi: 10.1037/0033-295X.110.4.683

Pecenka, N., & Keller, P. (2009). Auditory pitch imagery and its relationship to musical synchronization. *Annals of the New York Academy of Sciences, 1169*, 282-286. doi: 10.1111/j.1749-6632.2009.04785.x

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

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

Pfordresher, P. Q., & Mantell, J. T. (2012). Effects of altered auditory feedback across effector systems: Production of melodies by keyboard and singing. *Acta Psychologica, 139*, 166-177. doi: 10.1016/j.actpsy.2011.10.009

Pitt, M. A., & Crowder, R. G. (1992). The role of spectral and dynamic cues in imagery for musical timbre. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 728-738.

Pylyshyn, Z. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review, 88*, 16-45.

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

Rosenberg, H. S., & Trusheim, W. H. (1990). Creative transformations: How visual artists, musicians and dancers use mental imagery in their work. In J. E. Shorr (Ed.), *Imagery: Current perspectives* (pp. 55-75). New York: Plenum.

Rusconi, E., Kwan, B., Giordano, B. L., Umiiltà, C., & Butterworth, B. (2005). Spatial representation of pitch height: The SMARC effect. *Cognition, 99*, 113-129. doi: 10.1016/j.cognition.2005.01.004

Schubert, E. (2004). Modeling perceived emotion with continuous musical features. *Music Perception, 21*, 561-585.

Sloboda, J. A. (1985). Expressive skill in two pianists: Metrical communication in real and simulated performances. *Canadian Journal of Psychology, 39*, 273-293.

Takahashi, N., & Tsuzaki, M. (2008). Comparison of highly trained and less-trained pianists concerning utilization of auditory feedback. *Acoustical Science and Technology, 29*, 266-273. doi: 10.1250/ast.29.266

Thompson, W. L., Kosslyn, S. M., Hoffman, M. S., & Van der Kooij, K. (2008). Inspecting visual mental images: Can people "see" implicit properties as easily in imagery and perception? *Memory and Cognition, 36*, 1024-1032. doi: 10.3758/MC.36.5.1024

Wachowicz, F., Stevens, C. J., & Byron, T. P. (2011). Effects of balance cues and experience on serial recall of human movement. *Dance Research, 29*, 450-468.

Wiens, P. W. (2002). *Expressive conducting* [Audiovisual Material].

Williamson, V. J., Baddeley, A. D., & Hitch, G. J. (2010). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: Comparing phonological similarity and pitch proximity. *Memory and Cognition, 38*, 163-175. doi: 10.3758/MC.38.2.163

Wöllner, C., & Williamon, A. (2007). An exploratory study of the role of performance feedback and musical imagery in piano playing. *Research Studies in Music Education, 29*, 39-54. doi: 10.1177/1321103X07087567

Wood, J. N. (2007). Visual working memory for observed actions. *Journal of Experimental Psychology: General, 136*, 639-652.

Woody, R. (1999). The relationship between explicit planning and expressive performance of dynamic variations in an aural modeling task. *Journal of Research in Music Education, 47*, 331-342.

Wu, J., Yu, Z., Mai, X., Wei, J., & Luo, Y. (2011). Pitch and loudness information encoded in auditory imagery as revealed by event related potentials. *Psychophysiology, 48*, 415-419. doi: 10.1111/j.1469-8986.2010.01070.x