

## MUSICAL IMAGERY AND THE PLANNING OF DYNAMICS AND ARTICULATION DURING PERFORMANCE

---

LAURA BISHOP

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

FREYA BAILES

*University of Hull, Hull, United Kingdom*

ROGER T. DEAN

*University of Western Sydney, Penrith, Australia*

**MUSICIANS ANTICIPATE THE EFFECTS OF THEIR** actions during performance. Online musical imagery, or the consciously accessible anticipation of desired effects, may enable expressive performance when auditory feedback is disrupted and help guide performance when it is present. This study tested the hypotheses that imagery 1) can occur concurrently with normal performance, 2) is strongest when auditory feedback is absent but motor feedback is present, and 3) improves with increasing musical expertise. Auditory and motor feedback conditions were manipulated as pianists performed melodies expressively from notation. Dynamic and articulation markings were introduced into the score during performance and pianists indicated verbally whether the markings matched their expressive intentions while continuing to play their own interpretation. Expression was similar under auditory-motor (i.e., normal feedback) and motor-only (i.e., no auditory feedback) performance conditions, and verbal task performance suggested that imagery was stronger when auditory feedback was absent. Verbal task performance also improved with increasing expertise, suggesting a strengthening of online imagery.

*Received: December 21, 2011, accepted February 28, 2012.*

**Key words:** musical imagery, sensory feedback, expression, planning, musical expertise

---

**D**URING PERFORMANCE, MUSICIANS USE parameters such as pitch, timing, dynamics, and articulation to communicate their expressive interpretation of a piece (Juslin, 2000; Palmer, 1997). Internal representations of the action sequences

needed to perform a particular interpretation are constructed prior to their performance (Palmer & Pfordresher, 2003). This process of planning involves anticipating the effects of those actions (Keller, Dalla Bella, & Koch, 2010; Keller & Koch, 2008; Koch, Keller, & Prinz, 2004; Stock & Stock, 2004), a process skilled musicians report conceptualizing in terms of online musical imagery (Holmes, 2005; Rosenberg & Trusheim, 1990; Trusheim, 1993): the conscious experience of music during performance that is not a consequence of its production or perception. The ability to imagine a desired interpretation is said by some musicians to be integral to expressive music performance (Holmes, 2005; Rosenberg & Trusheim, 1990; Trusheim, 1993). Exploration of this idea in the laboratory has the potential to inform performers, students, and educators.

Musical imagery ability outside the performance context improves as a function of musical expertise, (Aleman, Nieuwenstein, Bocker, & de Haan, 2000; Janata & Paroo, 2006; Pecenka & Keller, 2009; Pitt & Crowder, 1992) and the range of planning during performance, or the amount of material accessible at a given time, increases (Drake & Palmer, 2000; Palmer & Drake, 1997; Palmer & Pfordresher, 2003). Expert musicians are characterized by their ability to be precise but flexible in their use of expression and better able to convey different interpretations than non-expert musicians (Palmer, 1997). The ability to use musical imagery in expressive performance planning might contribute to musicians' success at realizing their expressive intentions.

The rapid rate at which action sequences can be prepared during skilled music performance means that some aspects of planning are not accessible to performers' conscious awareness or control (Lashley, 1951). Other aspects are accessible to conscious awareness and can be retrospectively verbalized (Bangert, Schubert, & Fabian, 2009; Chaffin, Lisboa, Logan, & Begosh, 2010; Chaffin & Logan, 2006), but it is unclear to what extent these plans take the form of musical imagery. Potential alternate forms of conscious planning include verbal self-instructions about what is to be done (e.g., reminding oneself to get louder at a particular section or use a particular fingering) and forms of imagery that do not

involve the experience of music (e.g., mentally picturing the markings printed on a musical score).

Auditory imagery has been found to play a critical role in sensorimotor synchronization (Keller, 2012). Synchronizing with external sound sources, such as other musicians when playing in an ensemble, involves predicting when events will occur and coordinating movements to produce sounds that occur at the same time. In an fMRI experiment, people tapped in synchrony with tempo-changing pacing signals while completing a working memory task that varied across conditions in its difficulty. Increased working memory demands were found to impair prediction abilities and were associated with decreased activity in brain regions implicated in auditory imagery and attention. These findings suggest that imagery may be important for the temporal coordination of actions, but further research is needed to establish its contribution to expressive performance. Though performance could not proceed without planning of any sort, perhaps automatic motor planning, unconscious expectations, and non-imagery forms of conscious planning are sufficient. Also unclear is whether the contribution of musical imagery to performance planning changes with increasing musical expertise. The aims of the present research were to investigate whether musical imagery can be used in the planning of expressive dynamics and articulation during piano performance and to assess the relationship between online musical imagery ability and musical expertise. The presence and absence of auditory and motor feedback were manipulated to create conditions in which reliance on imagery was likely to differ.

#### DYNAMICS AND ARTICULATION IN EXPRESSIVE MUSIC PERFORMANCE

Mental rehearsal allows performers to test potential interpretations of a piece and analyze anticipated results without interference from auditory or motor feedback (Bailes & Bishop, 2012; Connolly & Williamon, 2004; Cowell, 1926; Holmes, 2005; Rosenberg & Trusheim, 1990; Trusheim, 1993). The use of mental rehearsal by many musicians in the Western classical tradition suggests that both note structures and the associated parameters of expression can be imagined. Musical expression is the systematic deviation from a prescribed tonal-temporal structure that constitutes an interpretation of a piece. Among other functions, expression reflects the performer's understanding of musical structure (Clarke, 1993; Juslin, 2003; Palmer, 1997). Generative models propose that expressive performance is guided by a cognitive representation of music structure that is translated into

modifications of acoustic parameters in accordance with style-specific conventions. These modifications emphasize information such as group boundaries, meter, and harmonic structure (Clarke, 1993; Juslin, 2003). Performers can manipulate parameters such as timing, dynamics, and articulation to convey their interpretation of the underlying musical structure (e.g., Juslin & Laukka, 2003).

Expressive timing has been investigated in a number of studies (Bangert et al., 2009; Clarke, 1993; Repp, 1997, 1999; Takahashi & Tsuzaki, 2008), but dynamics and articulation are also among the parameters that musicians most commonly manipulate (Juslin & Laukka, 2003), and they have received little attention in the literature (though see Kendall & Carterette, 1990; Nakamura, 1987; Palmer, 1989). People have been found to anticipate the loudness of sounds produced by their actions when making motor responses to visual cues (Kunde, Koch, & Hoffman, 2004). Whether performers anticipate dynamics, or *loudness change*, however, requires further study. Dynamics rather than the loudness of individual notes are typically used by performers to convey their expressive interpretations. Whether articulation is imagined during performance, also, has not been previously investigated. The present study, therefore, aimed to investigate the extent to which imagery for dynamics and articulation can contribute to expressive performance planning.

#### ONLINE MUSICAL IMAGERY AS A FORM OF CONSCIOUSLY ACCESSIBLE PLANNING

Skilled music performance is highly automatized (Duke, Cash, & Allen, 2011): action sequences are executed with minimal cost to attentional resources and with minimal interference to simultaneously occurring processes (Schneider & Shiffrin, 1977). While it has been posited that automatization of movements may enable performers to focus attention on higher-order processes, such as planning (Beilock, Wierenga, & Carr, 2002), not all planning is attentionally demanding either. Knowledge of musical structure, for example, can have a tacit influence on performance, constraining the time course of planning (Palmer & Pfordresher, 2003) and rendering tonally typical pitch errors more common than others (Palmer & Van de Sande, 1993).

Some planning may be consciously experienced and controlled by the performer, though, as observed in a case study of expert violin performance. Bangert et al. (2009) compared the number of musical decisions made with the performer's "conscious awareness" (i.e., verbalized or notated) to the number made without. Decisions

included instances of altered note duration, dynamics, and articulation, among others, and the vast majority observed were said to be made “without conscious awareness,” as they were perceptible during a final performance but not identified during practice. It was concluded that while the violinist’s expressive performance involved some aspects of planning that were deliberate, or accessible to conscious awareness and control, he also relied extensively on “musical intuition,” which occurs automatically. Further study is needed to assess the generalizability of these observations. Also, since only plans articulated in advance of the final performance were counted as deliberate, it is unclear whether the results provide an accurate depiction of the extent to which consciously accessible forms of planning are used.

Research on aural modelling, in contrast to the findings by Bangert et al. (2009), suggests that musicians who deliberately plan expressive parameters may be more likely to realize their plans during performance than musicians who do not (Woody, 1999, 2003). Aural modelling involves imitating a sounded performance and is a common method of music instruction (Dickey, 1991; Laukka, 2004; Lindström, Juslin, Bresin, & Williamon, 2003; Woody, 1999, 2003). Woody (1999) asked pianists to imitate performances of simple passages containing dynamic variations. Dynamics were more successfully replicated by pianists who were able to verbally identify them prior to performance than by pianists who did not verbally identify them, and it was concluded that performers who have a conscious intention to play specific patterns of dynamics are more likely to play them than performers who rely on musical intuition or summon up a particular feeling and trust that feeling to be encoded automatically in their playing. The nature of the explicit knowledge used by musicians in Woody (1999) is not clear, though. As the experimental passages were short, pianists may have encoded information about dynamics verbally (e.g., instructing themselves to “play a crescendo in the second bar”). In a normal performance context, when pieces can be long and expressive features numerous and overlapping, encoding all expressive information in this manner is unlikely to be effective or even feasible. If expressive information instead were represented in a guiding musical image, perhaps multiple expressive parameters could be integrated into the same performance plan with greater efficiency. In the present research, consistent with musicians’ self-reports, it was hypothesized that consciously accessible planning taking the form of musical imagery could be used during expressive performance.

#### THE ROLE OF SENSORY FEEDBACK IN MUSIC PERFORMANCE PLANNING

Performers can gauge how successfully they have realized their plans by monitoring the effects of their actions. Research on perceptual-motor coordination suggests that associations between actions and their auditory effects develop with experience (Keller & Koch, 2008), and that action sequence production is facilitated when actions elicit expected, rather than unexpected, acoustic effects (Keller et al., 2010; Keller & Koch, 2008; Kunde et al., 2004). Performance errors increase when auditory feedback is altered either tonally or temporally (Couchman, Beasley, & Pfordresher, 2012; Furuya & Soechting, 2010; Pfordresher, 2003; Pfordresher & Mantell, 2012), indicating that coupling between actions and acoustic effects enables fluent performance to be maintained.

Technical fluency does not depend on auditory feedback being present during the performance of learned pieces, however (Finney & Palmer, 2003). Some control over expression also remains in its absence (Repp, 1999, 2001; Takahashi & Tsuzaki, 2008; Wöllner & Williamon, 2007). Repp (1999) compared the expressive timing and dynamics of performances produced by skilled pianists in silence with those produced under normal auditory feedback conditions. Though the effects of auditory feedback deprivation were statistically significant for both parameters, they were so slight that listeners had difficulty distinguishing between performances produced with auditory feedback and performances produced in silence. It was concluded that auditory feedback may be used in the fine-tuning of expression, but that pianists can control expressive timing and dynamics meaningfully in the absence of sound.

Motor feedback seems to make a substantial contribution to performance planning as well. At a neural level, error monitoring in piano performance begins prior to errors being produced, whether or not auditory feedback is available (Ruiz, Jabusch, & Altenmüller, 2009). Erroneous pitches have been found to be played with reduced loudness, suggesting that the motor system includes a feed-forward system that regulates precision and accuracy in performance (Ruiz et al., 2009). Motor feedback begins to be available before a note is played, and as a result, may contribute to planning at least as much as auditory feedback, which is not received until after the execution of a corresponding action.

Musicians seem to use motor feedback to achieve fluency in performance. Keller et al. (2010) trained musicians to respond to visual stimuli (colors) by pressing vertically aligned buttons in a predefined pattern, at

a predefined pace. Each button press could elicit a high, medium or low-pitched tone. Timing was most accurate when auditory feedback was absent, but movement amplitude and acceleration were reduced when it was present. Exaggerating their movements likely increased the strength of the motor feedback musicians received, and strengthened motor feedback may have been needed to compensate for missing auditory feedback in musicians' attempts to maintain a steady tempo during the silent condition. It may be that the presence of motor feedback facilitates auditory imagery, as a result of the auditory-motor coupling that underlies action sequence production in a musical context (Hickok, Buchsbaum, Humphries, & Muftuler, 2003; Keller et al., 2010; Keller & Koch, 2008; Kunde et al., 2004).

Motor feedback contributes to musicians' success at achieving their expressive intentions as well. Wöllner and Williamon (2007) simultaneously disrupted auditory and motor feedback by asking skilled pianists to tap out the beat to an imagined performance in silence. They compared the resulting timing profiles to those produced at the piano during silent performances with normal motor feedback, and to those produced under normal auditory and motor conditions. The simultaneous disruption of both motor and auditory feedback affected performance more than the disruption of auditory feedback alone, and this was taken to be indicative of the importance of motor feedback to the realization of expressive intentions. It is unclear how comparable the performances produced under different feedback conditions in this experiment are, however, given the different methods used to disrupt the two types of feedback. Furthermore, it is possible that in tapping out a regularly occurring beat, pianists were encouraged to imagine performances "metronomically," and that this, rather than the disruption of motor feedback, led to a decline in expressivity. Further research is thus needed to investigate the contribution of motor feedback to expressive performance.

Musical imagery may guide performance when sensory feedback is missing or degraded (Repp, 2001). While motor programs can explain how simple musical sequences—or even complex sequences, given enough practice—can be played accurately and expressively without auditory feedback (Keele, 1968; Lashley, 1951; Schmidt, 1975), they do not explain why the performance of novel action sequences is disrupted when anticipated and perceived effects of actions do not match (e.g., Keller et al., 2010; Pfordresher, 2003). In the study by Keller et al. (2010), timing was more accurate when button-presses elicited tones that were compatible with participants' expectations in terms of spatial and pitch height

pairings (e.g., the top button elicited a high tone) than when they were incompatible (e.g., the top button elicited a low tone). These effects were said to derive from the use of anticipatory auditory imagery in action planning because movement acceleration was affected for the first tap of each sequence that was performed within a specific blocked condition, in response to a visual cue, before any tones for that sequence had sounded. Had the effects on timing been the result of automatic responses to unexpected auditory feedback, they would not have been observed until after auditory feedback had been received (e.g., Lashley, 1951). Furthermore, had participants relied on motor programs to execute the action sequences, then no effect of altered auditory feedback would have been predicted, as motor programs are said to be uninfluenced by auditory feedback. The results of the studies by Keller et al. (2010) and Wöllner and Williamon (2007) suggest that in silent performance conditions, auditory imagery may be stronger when normal motor feedback is present than when both auditory and motor feedback are disrupted. In the present study, the effects of auditory and motor feedback deprivation on the performance of dynamics and articulation were investigated, with different feedback conditions expected to yield different degrees of reliance on auditory imagery. It was hypothesized that stronger imagery for dynamics and articulation would be demonstrated during piano performance with normal motor feedback but no auditory feedback than during an entirely imagined performance, when both auditory and motor feedback were absent.

#### MUSICAL EXPERTISE AND WORKING MEMORY CAPACITY

Musical imagery engages working memory (Aleman & Wout, 2004; Kalakoski, 2001), and a possible explanation for any effects of expertise observed in imagery tasks is that expert musicians have a greater working memory capacity than novices. Some research suggests a relationship between musical expertise and working memory capacity, but the evidence is conflicting (Jakobson, Cuddy, & Kilgour, 2003; Wilson, 2002). In the present study, the working memory capacity of participants was assessed with a task tapping into the central executive to verify that this variable could not account for any differences in imagery task performance observed between expertise groups.

#### PRESENT RESEARCH

Many of the processes involved in producing music are not readily verbalized by performers (Beilock, Bertenthal, Hoerger, & Carr, 2008; Lindström et al., 2003), as is the case for any task in which automatization and implicit knowledge play a role. Skilled performers and educators

TABLE 1. Auditory and Motor Feedback Conditions.

Condition	Auditory feedback present	Motor feedback present	Stimulus presentation
Baseline performances	Yes	Yes	Unfolding score with <i>no</i> dynamic/articulation markings
Auditory-motor performance	Yes	Yes	Unfolding score with dynamic/articulation markings
Motor-only performance	No	Yes	Unfolding score with dynamic/articulation markings
Imagined performance	No	No	Unfolding score with dynamic/articulation markings
Perceptual	Yes	No	Unfolding score with dynamic/articulation markings

often explain how they go about creating an expressive performance using demonstrations (Dickey, 1991; Laukka, 2004; Lindström et al., 2003; Woody, 1999, 2003) or metaphors (Laukka, 2004; Lindström et al., 2003). Nevertheless, much of the research conducted on conscious planning has relied on case studies and verbal reports (Bangert et al., 2009; Chaffin et al., 2010; Chaffin & Logan, 2006). The current study, in contrast, used a piano performance task with different sensory feedback conditions to investigate how consciously accessible planning in the form of online musical imagery contributes to the performance of dynamics and articulation. The potential relationship between musical expertise and the use of online imagery in planning was also assessed.

Participants learned two simple melodies using notation devoid of expressive markings, then performed them expressively under three different feedback conditions: *auditory-motor* (i.e., normal auditory and motor feedback), *motor-only* (i.e., no auditory feedback, but with normal motor movements made on the keyboard), and *imagined* (no auditory or motor feedback) (Table 1). Dynamic and articulation markings were periodically introduced into the score during performance under these conditions, and a verbal compatibility judgement task required participants to indicate verbally (by saying “yes” or “no”) whether each marking matched or did not match their expressive intentions while continuing to play with their own interpretation. Participants gave a baseline performance under normal auditory and motor feedback conditions prior to completing the auditory-motor, motor-only, and imagined performance conditions, and this baseline performance was assumed to be indicative of their expressive intentions.

The strength of online imagery under different feedback conditions was assessed on the basis of participants’ success on the verbal compatibility judgement task. The timing and accuracy of verbal responses were together expected to indicate whether participants were imagining the dynamics and articulation that they intended to play. If imagery was used, correct verbal responses (i.e., responses that matched the dynamics

or articulation of the baseline performance) would be made prior to completing performance of the corresponding note sequence. If participants instead relied on the retrospective evaluation of auditory or motor feedback, then correct verbal responses would be made after performing the corresponding note sequence. A perceptual condition was included for comparison against the auditory-motor, motor-only, and imagined performance conditions, and to verify participants’ familiarity with and ability to perceive changes in dynamics and articulation. In the perceptual condition, participants listened to one of their own performances and completed the verbal compatibility judgement task based on what they were hearing. Working memory span was also assessed with an automated version of the Operation Span Task (OSPAN).

The contribution of auditory feedback to expressive performance was assessed by measuring the similarity between participants’ baseline performances and each of the auditory-motor and motor-only performances. It was hypothesized that for all expertise groups, baseline performances would be most accurately replicated when normal auditory and motor feedback were available (i.e., in the auditory-motor condition), given that performance conditions were most similar to the baseline in that case. However, imagery for dynamics and articulation was hypothesized to be strongest during the motor-only performance condition, when auditory feedback was absent but motor feedback was present. The lack of sound was expected to encourage the use of auditory imagery to guide performance (Brown & Palmer, 2012; Highben & Palmer, 2004; Repp, 1999), and it was thought that the presence of motor feedback might facilitate auditory imagery (Keller et al., 2010; Wöllner & Williamon, 2007). In line with previous research suggesting that motor feedback makes a substantial contribution to performance (Keller, et al., 2010; Wöllner & Williamon, 2007), during the imagined performances, the simultaneous absence of auditory and motor feedback was expected to lead to the weakest imagery for both expressive parameters. Working memory capacity was not expected to account for performance on the imagery task.

In most studies of musical expertise, comparisons are made between expert and novice performer groups. The current study included an intermediate group as well. Musical imagery was predicted to strengthen with increasing expertise, and the inclusion of novice, intermediate, and expert pianist groups meant that the difference in imagery ability between novice and intermediate groups could be compared to the difference in imagery ability between intermediate and expert groups.

## Method

### PARTICIPANTS

Twenty-nine pianists participated in the study, all recruited from universities and music schools in the Greater Sydney area. Three groups of pianists were invited to participate: experts, intermediates, and novices. Experts were required to either play the piano professionally or have a university degree in piano performance, and all reported between 10 and 20 years of private piano lessons. Intermediates were required to have had at least five years of private piano lessons (reported range 9-20 years), but not play the piano professionally and not have a university degree in piano performance. Novices were required to have between one and five years of private piano lessons (reported range 1-5 years), to not have a university degree in music, and to not play music professionally. Participants received either a small travel reimbursement or course credit.

Eleven experts completed the experiment (7 female, age  $M = 33.0$ , OMSI  $M = 625$ ), 11 intermediates (9 female, age  $M = 32.2$ , OMSI  $M = 462$ ), and seven novices (4 female, age  $M = 28.4$ , OMSI  $M = 319$ ). An additional nineteen people completed the experiment, but their data are not included in the analysis because they were either unable to meet the minimum technical skill requirements in their performances (14 participants; see Results for exclusion criteria), or, upon providing detailed information about their musical background, were found not to fit the criteria for any of the expertise groups (5 participants). Though the number of excluded participants is very large, it should be noted that nearly all of the participants who were unable to meet the technical requirements were undergraduate psychology students who, upon arriving for the experiment, reported not playing the piano regularly or recently and being unconfident with music reading. Many self-rated themselves as nonmusicians at the time of the experiment despite having had several years of formal piano lessons in the past.

### STIMULI

The right-hand lines from two pieces of Romantic-style piano music were selected for use in the experiment: Harthan's *Miniaturen*, Op. 17, No. 1, and Kabalevsky's *Pieces for Children*, Op. 39, No. 13 (Waltz). Both pieces came from collections written for beginning piano students and were selected because they were short, relatively unknown, and had technically simple but engaging melody lines carried by the right hand. They shared the same key signature, with one in F major (Harthan) and one in D minor (Kabalevsky). The piece by Harthan was in duple meter and 24 bars in length. The piece by Kabalevsky was in triple meter, and the first 32 bars were selected for use in the experiment. One expert participant reported some familiarity with the piece by Kabalevsky, but had not heard or performed it for many years. No other participants reported any familiarity with either piece.

Participants learned the pieces and gave initial baseline performances from scores devoid of expressive notation. Some fingering was given, but participants were free to follow it or not as they pleased. For use in the remaining performance and perceptual conditions, four versions of the score for each piece were developed. Each version contained a different arrangement of expressive markings. These markings indicated either articulation (staccato, slurs, breaks) or dynamics (crescendo, decrescendo, sforzando). All markings were symbols, not words, except for sforzando, which was represented by the customary *sf*. A total of six markings were introduced in each version: three indicated articulation and three indicated dynamics. Three of the markings in each version were taken from the original score, notated by the composer, and were therefore expected to align with a conventional expressive interpretation of the piece (hereafter referred to as "original" markings). Three other markings were added that were not in the original score and were expected to contradict a conventional expressive interpretation (hereafter referred to as "idiosyncratic" markings; Figure 1). It was acknowledged that performers could differ greatly in their expressive interpretation, and that as a result, original markings would not always be compatible with individual participants' intentions, and idiosyncratic markings would not always be incompatible with their intentions. However, it was still reasonable to assume that all or most participants would encounter markings that were both compatible and incompatible with their expressive intentions during each condition.

To introduce as much variety as possible into the four versions of each score and prevent participants from learning to expect particular markings or markings in

Figure 1 consists of two parts, (a) and (b), each showing two lines of musical notation. Part (a) shows a two-line extract from Harthan's *Miniaturen*, Op. 17, No. 1. The notation includes various note values, rests, and expressive markings such as a crescendo hairpin and a final slur. Part (b) shows a two-line extract from Kabalevsky's *Pieces for Children*, Op. 39, No. 13 (Waltz). This notation includes dynamic markings (piano and forte), slurs, and other expressive notations. The musical notation is presented in a standard staff format with a treble clef and a key signature of one flat.

FIGURE 1. Scores for experimental stimuli with sample expressive notation. (a) Two line extract from Harthan's *Miniaturen*, Op. 17, No. 1. Only the crescendo and final slur were expected to be compatible with a conventional interpretation. Pianists were instructed to play at one quarter note = 108 bpm. (b) Two line extract from Kabalevsky's *Pieces for Children*, Op. 39, No. 13 (Waltz). The first and last dynamic markings and the first slur were expected to be compatible with a conventional interpretation. Pianists were instructed to play at one quarter note = 138 bpm.

particular locations, seven (Kabalevsky) or eight (Harthan) locations were selected for each piece and markings were displayed at six of them in each trial. These locations corresponded to a variety of structural features, including key modulation, repeated phrases, phrase boundaries (as indicated in the original score), and the final phrase of the piece. There were never any markings among the first five notes and consecutive markings were always separated by at least four beats. Though an original marking occasionally appeared in the same location in two different scores, idiosyncratic markings did not, in case their distinctiveness made them more memorable.

Scores were displayed to participants on a computer screen via PowerPoint. The entire score for each piece

(without expressive markings) was available to participants as they practiced. During the initial baseline performance and subsequent experimental conditions, however, music was revealed gradually as the participant progressed through the piece. Scores were revealed gradually to prevent participants from scanning ahead at the start of the performance and forming judgements about expressive markings earlier than they would under normal performance conditions, particularly during the imagined performance. It is acknowledged that some participants would adopt such a strategy during normal performance, but in this experiment, the focus was on the shorter-range planning that occurs immediately prior to notes being performed. At the

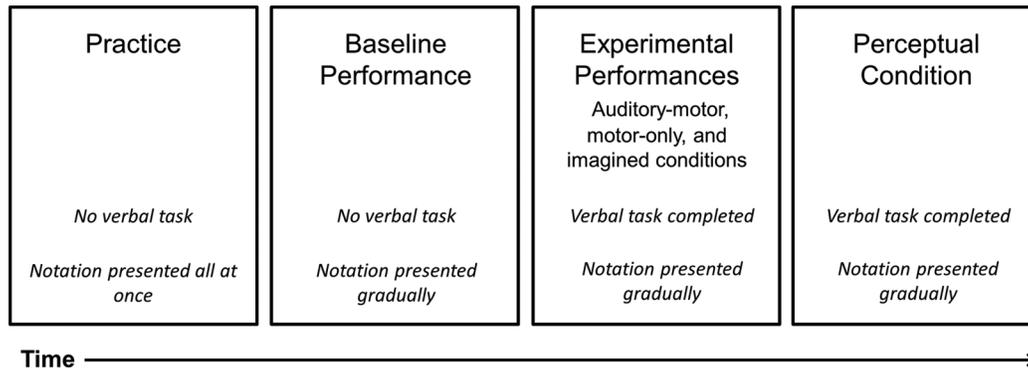


FIGURE 2. Schematic of experimental procedure.

start of each trial, the first five notes of the piece were presented on a staff that was otherwise blank apart from barlines. The remainder of each score appeared one bar at a time, dissolving in across a 0.2 s period. Once a bar was displayed on the score, it remained visible until the end of the performance. Participants were asked to keep more or less to a specific tempo for each piece (108 bpm for Harthan; 138 bpm for Kabalevsky), selected from within the ranges specified on the original scores. Pianists were reminded of this tempo with two bars of sounded metronome beats prior to starting each new performance. The pace at which music was revealed on the slides corresponded to this selected tempo. With participants playing at approximately the desired pace, at least the four subsequent beats were available for reading.

#### EQUIPMENT

All data were collected in the Virtual Interactive Performance Research Environment (VIPRE) at the University of Western Sydney. Participants were seated at an acoustic grand piano, the Yamaha Disklavier 3 (which has MIDI sensors and keys that can be MIDI-activated, in a performance studio, wearing an AKG C417 lapel microphone and AKG K271 Studio headphones, through which they heard sampled piano sound from the disklavier. The experiment was controlled from two MacBooks (OS X 10.5.8 and OS X 10.4.8) in an adjacent room. The participant and experimenter could see each other through a large window and microphones enabled verbal communication between them.

One of the MacBooks was connected to a monitor, which was placed at the participant's eye level on top of the closed lid of the piano. Music scores were presented to participants on this monitor using Microsoft PowerPoint. Each score was displayed on one PowerPoint slide containing staves one cm in height and 26.3 cm in

length. Audio from the performer's microphone and MIDI data from the piano were collected on the second MacBook, using a custom-designed patch in Max/MSP.

An automated version of the Operation Span Task (OSPAN) was presented to participants on a PC with Inquisit.

#### DESIGN

Performers completed the auditory-motor, motor-only, and imagined conditions in a random order, followed by the perceptual condition. The version of the scores performers saw during each performance condition and the perceptual condition was counterbalanced, with the two performers within an expertise group who completed the conditions in the same order being assigned different versions of the scores for each condition. Across all groups, half the participants began each condition with the piece by Harthan and half began with the piece by Kabalevsky.

#### PROCEDURE

Figure 2 illustrates the order in which different parts of the experiment were completed by each participant. At the start of the experimental session, participants received written and verbal instructions and completed a musical background questionnaire (including all questions from the Ollen Musical Sophistication Index (OMSI); Ollen, 2006). They were then given up to 30 min to practice the two pieces. The same MIDI grand piano was used for both the practice and performance segments of the session. Participants were instructed to practice the pieces until they could play them expressively, as though for a music lesson. They were not given the name or composer of either piece and never saw the left hand lines. They were told that some fingering was given for their reference, but that they were free to follow it or not as they chose.

Before participants began practicing a piece, they listened to a metronomic rendition of it played to them through the piano. They could see the keys moving and heard the sound through their headphones. These metronomic performances were generated using a custom-designed patch in Max/MSP and contained all the correct pitches and note durations, presented with a constant key velocity and at the target tempo. Participants were asked to “give a more interesting and expressive performance of each piece at this tempo.” They were free to spend as much of the 30 min practicing as they felt necessary, and when either the allotted practice time had expired or they decided they were ready, the performance segment of the session began.

All performances, as well as practice, were conducted while the participant was alone in the room. After the instructions were explained and the participant was given the opportunity to ask questions, the experimenter shut the door and went into an adjoining room. The experiment was run and data were fed into a computer controlled by the experimenter in this room.

Participants were reminded prior to giving performances that the focus of the study was expression, and they were asked to try to maintain a constant interpretation across all their performances of a piece. In order to establish baseline measures of expressive dynamics and articulation, participants first gave three performances of each piece while reading from scores devoid of expressive notation. They then indicated which performance they thought was their best, and MIDI data from that performance were used as the baseline reference series during analysis. The first five notes of a piece were visible on the score at the start of each trial, and two measures of beats were sounded by a metronome through a speaker in the room at the specified tempo. Participants were free to begin playing any time after this. Notes were revealed on the score automatically, one bar at a time, beginning at performance of the first note. The experimenter initiated this process each time a participant played a starting note.

After three baseline performances, the experiment proceeded with auditory-motor, motor-only, and imagined performance conditions, carried out in a counterbalanced order across participants. Participants performed each piece from a version of the score to which six expressive markings had been added, and music was revealed a bar at a time on the computer screen, as during the baseline performances. Tempo, again, was set prior to each performance. Whenever they came to an expressive marking in the score, participants were to continue playing with their own practiced interpretation and verbally judge the compatibility of the marking with their expressive

intentions for that segment of music by saying “yes” if the marking was compatible with their intentions and “no” if it was not, without interrupting their performance. They were told to focus on playing with their practiced interpretation and avoid introducing new expressive features to their performances, including those displayed in the score. Participants wore a microphone and their verbal responses were recorded.

Each condition began with a practice trial. Participants were presented with the score for an ascending C major scale, written as eight quarter-notes (crotchets) on a single staff, and asked to play it at a specified tempo and with a crescendo. A dynamic marking appeared when the second bar was revealed, and participants were to say “yes” if this marking matched their intended crescendo and “no” if it did not. This practice trial was done under the same feedback conditions as the performance the participant was about to undertake.

During the auditory-motor performance, participants could hear themselves as normal, receiving immediate, unaltered auditory feedback through headphones. During the motor-only performance, the headphones were disconnected and participants played the pieces in the absence of auditory feedback. During the imagined performance, participants did not play the pieces, but instead were asked to sit at the piano and imagine playing them while completing the verbal compatibility judgement task. They were told not to move their fingers, and to place their hands on top of the piano so that the experimenter could check for compliance.

The perceptual condition was always completed last. Participants were told that they would hear a participant’s performance of each piece and were asked to judge the compatibility of the expressive markings displayed on the score with the music they were hearing. They listened to their self-chosen ‘best’ baseline performances (though they were not informed that it was their own) while completing the verbal compatibility judgement task. The sounded performance was played to them by the piano so that they saw the keys moving and heard the music through their headphones.

*Automated Operation Span Task (OSPAN).* Working memory was evaluated with an automated version of the Operation Span Task created by Turner and Engle (1989). This task is high in validity and reliability (Conway et al., 2005) and has been used to measure working memory capacity in our previous research (Bishop, Bailes, & Dean, 2013). During the task, a simple mathematical equation was displayed on a computer screen (e.g.,  $(3 \times 2) + 1 = ?$ ), followed by a number. Participants indicated whether this number was the correct answer for the equation. They were then shown a letter, and

TABLE 2. Mean Total Pitch and Timing Errors Across Conditions.

Condition	All	Expertise group		
		Novice	Intermediate	Expert
Baseline	1.0 (1.2)	1.7 (1.0)	1.0 (1.6)	0.5 (0.7)
Auditory-motor	1.9 (2.5)	4.1 (3.4)	1.6 (2.2)	0.7 (0.9)
Motor-only	1.8 (1.9)	3.5 (2.0)	1.5 (1.7)	0.9 (1.2)

Note. These values represent only the errors made by the participants included in the rest of the analyses ( $n = 29$ ), and exclude those who did not meet the technical requirements or fit the criteria for any of the expertise groups. Standard deviation is in parentheses.

after solving between two and seven equations, had to recall the letters in the order that they were presented. Participants were asked to respond as quickly as possible and to maintain a minimum accuracy rate of 85% on the math. Instructions and practice trials were presented via the computer.

At the end of the session, participants were fully debriefed with regards to the aims and hypotheses of the study and asked to respond to a short questionnaire about their experience in completing the experiment.

## Results

### ERRORS IN PERFORMANCE FLUENCY

Performances were analyzed only if pitch or timing errors occurred on fewer than 15% of notes. Timing errors were said to occur when a note duration fell outside a specified range, which was established to allow for expressive timing and some between-subject differences in articulation and tempo. To calculate this range, interonset interval (IOI) profiles for eight randomly selected participants' baseline performances were averaged across participants to obtain a mean IOI for each note event in a piece. A mean IOI for each note category (e.g., quarter note, half note, etc.) across performances of a piece was also calculated, based on IOI profiles for the same eight participants. Timing errors were defined as note durations that fell (1) more than three standard deviations away from the mean for a particular note event and (2) within the mean range for a different note category. This method of identifying timing errors did not permit substantial deviations from the prescribed overall tempo, long hesitations, or inaccurately performed rhythms, but accepted moderate variations in note durations resulting from expressive timing and articulation. All data from 14 participants (7 novices, 7 intermediates) were excluded because the minimum accuracy requirement was not reached for either piece. One additional novice's auditory-motor performance of the Kabalevsky piece was excluded on this basis as well (total 57 trials excluded).

Among participants who met the inclusion criteria, fewer pitch and timing errors were made in baseline performances than in either the auditory-motor or motor-only performances, reflecting the added demands of the verbal task. There was no significant difference between error rates in auditory-motor and motor-only performances with data pooled across the two pieces, however,  $t(57) = 0.38$ ,  $p > .05$  (Table 2). Errors declined with increasing expertise, which a MANOVA for each piece, in which the dependent variables were total errors in baseline, auditory-motor, and motor-only conditions and the independent variable was expertise group, was significant for Wilks' lambda,  $F(2, 28) = 2.89$ ,  $p < .05$  (Harthan),  $F(2, 28) = 2.48$ ,  $p < .05$  (Kabalevsky).

### WITHIN-SUBJECT CONSISTENCY IN DYNAMICS AND ARTICULATION

It was predicted that auditory-motor performances would be more similar than motor-only performances to baseline performances in terms of dynamics and articulation. To assess the similarity between performances, dynamic and articulation profiles for each participant's performances were first extracted. Dynamic profiles comprised the series of key velocities, and articulation profiles comprised the series of note duration to IOI ratios (see Friberg & Battel, 2002; Hähnel & Berndt, 2010). The similarity between dynamic profiles and articulation profiles produced under different feedback conditions was then measured using dynamic time warping (DTW) (Giorgino, 2009), which is suitable for use with time series data as it does not require independence of data points within participant profiles.

DTW identifies points along two data profiles (a 'reference' profile and 'test' profile) that likely correspond to each other, and calculates an average distance between profiles per note, with no penalty for missing data points (e.g., notes skipped or removed due to error). In the present analyzes, each participant's auditory-motor and motor-only performance (test) profiles were aligned with their baseline (reference)

TABLE 3. Mean DTW Distance From Baseline.

Condition	Expertise group			
	All	Novice	Intermediate	Expert
<i>Dynamics</i>				
Auditory-motor	3.18 (.95)	3.49 (.84)	3.29 (.92)	2.87 (1.01)
Motor-only	3.49 (.98)	3.62 (.72)	3.50 (.90)	3.38 (1.21)
<i>Articulation</i>				
Auditory-motor	0.06 (.04)	0.06 (.03)	0.07 (.04)	0.05 (.04)
Motor-only	0.07 (.04)	0.07 (.04)	0.07 (.04)	0.06 (.03)

Note. Three outliers have been removed from the distribution of DTW distances for dynamics and two outliers have been removed from the distribution of DTW distances for articulation. Standard deviation is in parentheses.

profiles on the basis of the prescribed pitch, with empty data points left wherever a note was skipped or removed due to error. A symmetric step-pattern was then used to calculate the average distance between the series of key velocities or articulation ratios that corresponded to the series of correctly performed pitches for baseline, auditory-motor, and motor-only conditions. Step-patterns dictate which points on a test profile can be matched to each point on the reference profile; symmetric step-patterns allow each point on the test profile to be matched to multiple points on the reference profile. This was necessary as errors in performance meant that participant profiles had occasional missing data points. Mean distances achieved by novice, intermediate, and expert pianist groups under auditory-motor and motor-only conditions are given by absolute rather than signed values and are reported in the same units as used by the test and reference data profiles (Table 3).

To investigate the potential effects of auditory feedback deprivation and musical expertise on the reliability of performed dynamics and articulation, DTW distances were entered into an ANOVA for each parameter. There were two independent variables for each ANOVA (feedback condition, expertise group), and DTW distance acted as the dependent variable. Data from both pieces were included. Three outliers were removed from the distribution of DTW distances for dynamics (two experts, one intermediate) and two outliers were removed from the distribution of DTW distances for articulation (both intermediates) because they fell more than 2.5 standard deviations from the mean. Since the distributions of DTW distances were highly skewed for both dynamics and articulation, log transformations were then conducted on each dependent variable to approximate normality. For dynamics, there was no significant main effect of feedback condition,  $F(1, 112) = 2.54$ ,  $p > .05$ , and no significant interaction between feedback condition and expertise,  $F(2, 112) = 0.38$ ,

$p > .05$ , but the main effect of expertise approached significance,  $F(2, 112) = 2.89$ ,  $p < .06$ . For articulation, there was no significant main effect of feedback condition,  $F(1, 112) = 0.35$ ,  $p < .05$ , and no significant interaction,  $F(2, 112) = 0.53$ ,  $p > .05$ , but the main effect of expertise was significant,  $F(2, 112) = 4.36$ ,  $p < .02$ . These findings suggest that success at replicating baseline dynamics and articulation improved with increasing expertise, as hypothesized, but contrary to expectations, pianists did not replicate their expressive baseline performance any more precisely under normal feedback conditions than in the absence of auditory feedback. Furthermore, the lack of interaction between feedback condition and expertise suggests that novice, intermediate, and expert pianists were similarly unaffected by auditory feedback deprivation.

The absence of a significant effect of auditory feedback deprivation might have resulted from participants performing inexpressively during the experimental performances. An additional post-hoc analysis was thus conducted to test whether dynamics and articulation had been performed as intended, at least to some degree, during the auditory-motor and motor-only conditions. DTW distances were calculated between the dynamic and articulation profiles for an expressively 'flat' performance and each participant's auditory-motor and motor-only performances. First, an overall mean key velocity and an overall mean articulation value for each piece was calculated by averaging the key velocities and articulation values performed by all participants in all feedback conditions. The expressively flat performance had a constant key velocity equivalent to this overall mean key velocity, and constant articulation equivalent to the overall mean articulation value. If baseline dynamics and articulation were maintained to some degree during the auditory-motor and motor-only performances, then participants' auditory-motor and motor-only performances should be more similar to

their baseline performances than to the expressively flat performance. Three outliers were excluded from the distribution of DTW distances separating auditory-motor and flat performances and three were excluded from the distribution of DTW distances separating motor-only and flat performances because they fell more than 2.5 standard deviations from the mean (in both cases, one intermediate for dynamics, one expert for dynamics, one intermediate for articulation). The five outliers identified in the earlier analysis were also excluded.

As predicted, Wilcoxon signed-rank tests (nonparametric alternative to *t*-tests to account for the skewed distributions of DTW distances) using data from both pieces showed that the mean DTW distance between participants' auditory-motor and baseline performances (dynamics  $M = 3.18$ ,  $SD = .95$ ; articulation  $M = 0.06$ ,  $SD = 0.04$ ) was smaller than the mean DTW distance between participants' auditory-motor performances and the expressively flat performance for both dynamics,  $z(54) = 5.83$ ,  $p < .001$  ( $M = 4.96$ ,  $SD = 1.85$ ), and articulation,  $z(56) = 4.70$ ,  $p < .001$  ( $M = 0.08$ ,  $SD = 0.04$ ). Similarly, the mean DTW distance between participants' motor-only and baseline performances (dynamics  $M = 3.49$ ,  $SD = 0.98$ ; articulation  $M = 0.07$ ,  $SD = 0.04$ ) was smaller than the mean DTW distance between participants' motor-only performances and the expressively flat performance for both dynamics,  $z(54) = 3.80$ ,  $p < .001$  ( $M = 4.48$ ,  $SD = 1.61$ ), and articulation,  $z(56) = 4.36$ ,  $p < .001$  ( $M = 0.08$ ,  $SD = 0.04$ ). These results suggest that pianists performed dynamics and articulation as intended, at least to some degree, during the auditory-motor and motor-only conditions. The lack of significant effects of auditory feedback deprivation were not a result of participants performing more inexpressively during the experimental performances than during the baseline performances.

#### VERBAL COMPATIBILITY JUDGEMENT TASK

*Verbal task accuracy.* The absence of sound during performance has previously been found to encourage the use of auditory imagery (Brown & Palmer, 2012; Highben & Palmer, 2004; Repp, 1999). It has also been suggested that the presence of motor feedback may facilitate auditory imagery, rendering auditory imagery during performance with motor feedback stronger than auditory imagery during an entirely imagined performance (Keller et al., 2010; Wöllner & Williamon, 2007). In the current study, it was therefore predicted that dynamics and articulation would be imagined during all performance conditions (auditory-motor, motor-only, and imagined) but be strongest during the

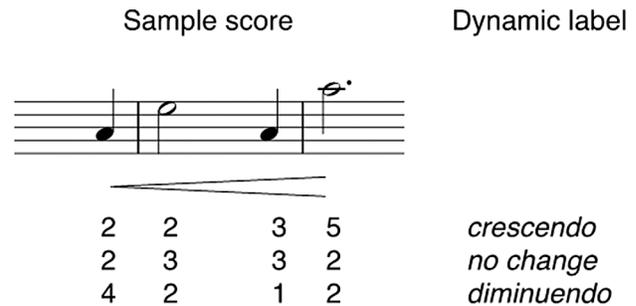


FIGURE 3. Illustration of dynamics coding process. Dynamic labels were assigned based on the key velocity gradient at the location of each dynamic marking. Key velocity profiles were divided into five quantiles for each performance, ranging from soft (quantile 1) to loud (quantile 5).

motor-only performance, when auditory feedback was absent but motor feedback was present. Verbal responses, accordingly, were expected to be fastest and most accurate during this condition. Participants were instructed to judge verbally whether each marking matched their expressive intentions, so assessing the accuracy of their responses first required determining whether they had played in a way that was consistent with those markings during their baseline performance, which was taken to be indicative of their intentions.

Performed dynamics at the location of each marking were categorized as *crescendi*, *diminuendi*, or *constant* by dividing key velocity distributions into five quantiles, and then assessing the gradient of these quantile values for the notes performed at the location of each marking (see Figure 3 for an illustration). Categories were within-subject, defined independently for each individual performance. Quantiles, instead of raw MIDI key velocities, were used so that the gradients identified would correspond to ranges of key velocities that were large enough to be perceptible. An ascending gradient indicated a *crescendo*, a descending gradient indicated a *diminuendo*, and a flat gradient indicated no measurable change in dynamics. A verbal response was said to be correct if the participant endorsed a marking that was performed in the baseline condition or correctly rejected a marking that was not.

To assess articulation, it was assumed that performance articulation could be divided into three categories, corresponding roughly to *legato*, *semi-detached*, and *staccato*. Tertiles were therefore calculated for the distribution of articulation values taken from all of eight randomly selected participants' baseline performances. This between-subjects distribution was used instead of identifying quantiles for each individual performance, as was done for dynamics, because it was anticipated that many

participants would not interpret the pieces as necessitating staccato and would only use two of the possible articulation categories. Accuracy of verbal responses to articulation markings was assessed on a note-by-note basis. If a participant endorsed a marking that consisted of a four-note slur, for example, a point was given for any of the first three notes that were performed with a legato or semi-detached articulation. Because the end of a slur indicates a contrast in the degree of connectedness between notes, a fourth point would be given if the final note had either a semi-detached or staccato articulation. Conversely, if the participant rejected a marking that consisted of four staccatos, four points were given unless all the notes had been performed as staccato in the baseline.

To test whether verbal task performance differed from chance, *t*-tests were conducted to compare the accuracy (i.e., proportion correct out of total responses) achieved in each condition with chance performance (50%). Verbal task accuracy was found to be significantly better than chance for the auditory-motor performance,  $t(55) = 10.51, p < .001$ , motor-only performance,  $t(49) = 10.59, p < .001$ , imagined performance,  $t(53) = 9.46, p < .001$ , and perceptual conditions,  $t(55) = 9.61, p < .001$ .

Contrary to expectations, the proportion of correct verbal responses did not differ either between conditions or between expertise groups: an ANOVA was conducted using verbal task accuracy as the dependent variable and condition and expertise group as independent variables, and this showed no significant main effect of condition,  $F(2, 159) = 0.50, p > .05$ , or group,  $F(2, 159) = 1.00, p > .05$ , and no significant interaction

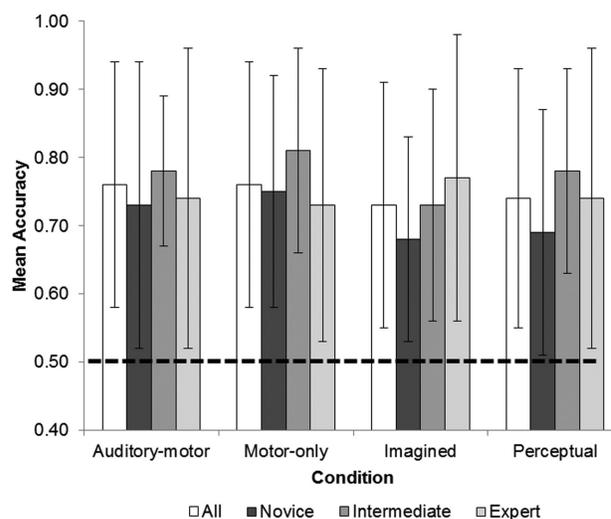


FIGURE 4. Mean accuracy for verbal compatibility judgement task. Accuracy indicates the proportion of correct answers out of total responses. The dotted line represents chance performance (50%) and error bars indicate standard deviation.

between them,  $F(4, 159) = 0.68, p > .05$ . Accuracy was similar even between performance and perceptual conditions (Figure 4).

Pianists in all expertise groups were significantly more likely to erroneously endorse markings taken from the original scores ('original' markings) than to endorse markings designed to contradict those taken from the original scores ('idiosyncratic' markings) (Table 4). This was established on the basis of a post-hoc analysis that investigated whether the false positive rates for original markings and the false positive rates for

TABLE 4. Proportions of Verbal Compatibility Judgement Task Errors Endorsing 'Original' and 'Idiosyncratic' Markings.

Condition	All	Expertise group		
		Novice	Intermediate	Expert
<b>Auditory-motor<sup>a</sup></b>				
Original	.62 (.38)	.46 (.35)	.67 (.36)	.60 (.43)
Idiosyncratic	.18 (.20)	.13 (.28)	.15 (.22)	.27 (.38)
<b>Motor-only<sup>a, c</sup></b>				
Original	.63 (.35)	.58 (.37)	.59 (.38)	.75 (.26)
Idiosyncratic	.11 (.21)	.15 (.21)	.18 (.27)	.01 (.03)
<b>Imagined<sup>a, b</sup></b>				
Original	.63 (.35)	.75 (.26)	.61 (.32)	.76 (.35)
Idiosyncratic	.11 (.22)	.14 (.20)	.17 (.28)	.04 (.13)
<b>Perceptual<sup>a</sup></b>				
Original	.55 (.39)	.58 (.39)	.64 (.37)	.46 (.37)
Idiosyncratic	.15 (.28)	.17 (.25)	.11 (.23)	.14 (.29)

Note. Standard deviation is in parentheses. (a) Significant effect of marking type,  $F(3, 105) = 23.48, p < .001$  (auditory-motor),  $F(1, 101) = 35.13, p < .001$  (motor-only),  $F(1, 103) = 12.59, p < .001$  (imagined),  $F(1, 101) = 8.67, p < .001$  (perceptual). (b) Significant interaction between marking type and expertise group,  $F(3, 103) = 2.78, p < .05$ . (c) Near-significant interaction between marking type and expertise group,  $F(3, 101) = 2.57, p < .06$ .

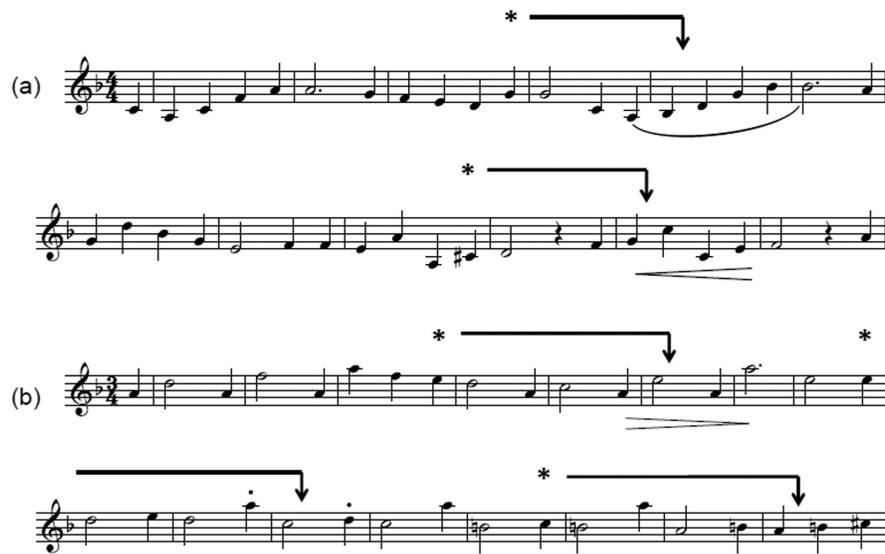


FIGURE 5. Verbal compatibility judgement task response times relative to scores. The first two lines from a version of the pieces by (a) Harthan and (b) Kabalevsky are shown; (\*) indicates location of performers at the time the next marking was revealed, assuming performance at approximately the instructed tempo, and arrows indicate performers' location at the time verbal responses were given (on average) during the motor-only performance.

idiosyncratic markings differed within each condition and between expertise groups. For each condition, an ANOVA was conducted using the proportion of errors that indicated false positive responses as the dependent variable and the type of marking (original or idiosyncratic) and expertise group as independent variables. Significant and near-significant results of these post-hoc tests are listed in Table 4 so as not to distract from the main results of the experiment. In all conditions, participants erroneously endorsed original markings significantly more often than they endorsed idiosyncratic markings. In motor-only and imagined conditions, the effect was particularly apparent for experts. These findings show that participants were biased in their verbal responses by how predictable or appropriate markings were in the context of a conventional interpretation of each piece. Though there is a significant interaction between marking type and expertise group in the imagined performance condition and the suggestion of an interaction in the motor-only performance condition, participants in all expertise groups were much more likely to erroneously endorse original markings than idiosyncratic markings, suggesting that novices, intermediates, and experts were all influenced in their verbal responses by schematic expectations relating to the predictability of the markings.

*Verbal task response times.* Verbal response times were coded manually using the voice and piano audio

recordings made during the experimental sessions. Response time was calculated as the distance in beats between the note being played when a marking appeared on the score and the note being played when the verbal response was produced (see Figure 5 for an illustration). Audio from the microphone worn by participants and audio from the piano were recorded on separate channels, so sound from the piano did not degrade voice recordings. Beats were used as the unit of measurement rather than seconds, since coding was done by ear and was therefore not precise at the millisecond level.

The speed of verbal responses differed significantly between conditions and expertise groups (Figure 6). A two-way ANOVA, in which the independent variables were feedback condition and expertise group and the dependent variable was response time in beats, was conducted to investigate the effects of auditory and motor feedback deprivation and expertise on the strength of imagery. Verbal response data for both pieces were included. This revealed a significant main effect of condition,  $F(5, 216) = 15.13, p < .001$ , and a significant main effect of expertise group,  $F(2, 216) = 6.62, p < .01$ , but no interaction between them,  $F(6, 205) = 0.48, p > .05$ . Planned comparisons indicated faster response times for all performance conditions than for the perceptual condition after a Bonferroni correction was applied,  $F(1, 205) = 27.73, p < .001$  (auditory-motor)  $F(1, 205) = 38.01, p < .001$  (motor-only),  $F(1, 205) = 22.84, p < .001$  (imagined). The faster response

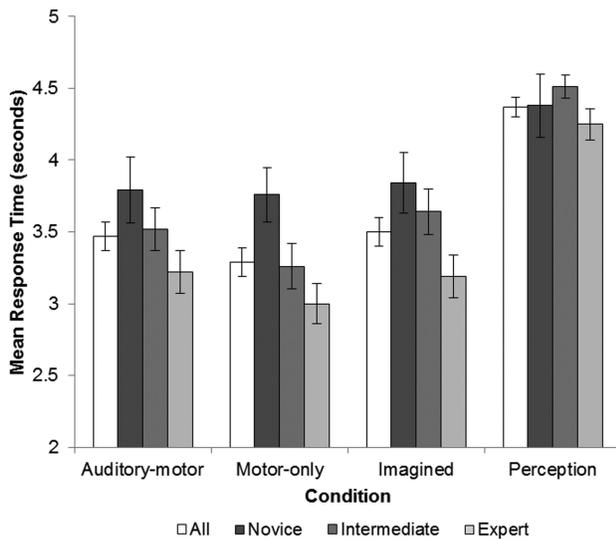


FIGURE 6. Verbal compatibility judgement task mean response times (seconds). Data for both pieces are included. Error bars indicate standard error.

times observed during the performance conditions, compared to the perceptual condition, suggest that pianists were responding in anticipation of the note sequences during the performance conditions, rather than retrospectively, or on the basis of auditory or motor feedback. Paired *t*-tests, again including data from both pieces and with a Bonferroni correction, indicated that responses were significantly faster during the motor-only performance than during either the auditory-motor,  $t(57) = 3.00, p < .01$ , or imagined performances,  $t(53) = 3.92, p < .001$ . This finding is in line with the hypothesis that imagery for dynamics and articulation is strongest when auditory feedback is absent but motor feedback is present: performance plans may have been most accessible to conscious awareness or specified information further in advance during the motor-only performance.

To check whether the faster response times observed for motor-only performances resulted from pianists playing faster than the prescribed tempo during that condition, baseline, auditory-motor, and motor-only performance durations were compared. An ANOVA, in which condition was the independent variable and performance duration was the dependent variable, showed no significant effect of condition for either piece,  $F(2, 95) = 1.03, p > .05$  (Harthan),  $F(2, 94) = 0.04, p > .05$  (Kabalevsky). This equivalence in performance durations across conditions suggests that differences in verbal task response times cannot be attributed to differences in performance tempo.

Response times were also found to improve with increasing expertise across conditions. Planned comparisons also showed experts' response times to be significantly faster than both novices' at a Bonferroni-adjusted alpha of .02,  $F(1, 205) = 12.05, p < .001$ , and intermediates' across conditions,  $F(1, 205) = 5.98, p < .02$ , though novices' and intermediates' response times did not differ from each other,  $F(1, 205) = 1.62, p > .02$ . These findings suggest that the strength of imagery improved with increasing expertise. Experts' plans may have been more accessible to conscious awareness than novices' or intermediates', or experts may have planned further ahead.

#### WORKING MEMORY CAPACITY

Working memory capacity, assessed using the OSPAN, was not expected to account for either the consistency of performed dynamics and articulation across conditions or performance on the verbal compatibility judgement task. An ANOVA was conducted using OSPAN scores as the dependent variable and expertise group as the independent variable, and OSPAN scores were not found to differ significantly between expertise groups. Novices achieved a mean score of 49 ( $SD = 18$ ), intermediates a mean score of 51 ( $SD = 16$ ), and experts a mean score of 46 ( $SD = 18$ ). Pearson's correlations were calculated between OSPAN score and DTW distances (subjected to a log transformation to approximate normality), verbal task response accuracy, and verbal task response time, but none was significant.

#### Discussion

The possibility that enhanced online imagery ability contributes to the high precision with which expert musicians realize their expressive intentions was addressed with an investigation of whether online imagery can contribute to the performance of dynamics and articulation and whether online imagery ability covaries with musical expertise. Novice, intermediate, and expert pianists developed an interpretation of two pieces using scores devoid of expressive markings, then performed those pieces while verbally judging whether dynamic and articulation markings periodically introduced into the score matched their expressive intentions. This was done under different sensory feedback conditions that were expected to encourage reliance on imagery to varying degrees.

Dynamics and articulation of the chosen interpretation were maintained during both the performance with auditory feedback (auditory-motor condition) and the performance without auditory feedback (motor-only

condition). No significant effect of auditory feedback deprivation on either parameter was found. Pianists' success at performing in the motor-only condition is consistent with prior research showing that music learned with auditory feedback can subsequently be performed with technical accuracy in the absence of auditory feedback (Brown & Palmer, 2012; Finney & Palmer, 2003; Highben & Palmer, 2004; Repp, 1999; Takahashi & Tsuzaki, 2008; Wöllner & Williamon, 2007). Pianists' success at maintaining dynamics and articulation in the motor-only condition, however, is at odds with previous research that has shown dynamics and expressive timing to be less reliable in the absence of auditory feedback than under normal feedback conditions (Repp, 1999; Takahashi & Tsuzaki, 2008). A potential explanation is that the added demands of the verbal task led participants to play inexpressively under both feedback conditions; however, post-hoc analyses revealed that participants' auditory-motor and motor-only performances were more similar to their baseline performances than to an expressively flat performance with constant key velocity and articulation, suggesting that the lack of significant effects was not a result of inexpressive playing during the auditory-motor and motor-only conditions. Instead, auditory feedback may not have been essential for participants to perform dynamics and articulation as intended in the context of the current study.

Analysis of responses made on the verbal compatibility judgement task showed that significantly above-chance rates of accuracy were achieved in all conditions, both perceptual and performance. This accuracy, coupled with the significantly faster average response times that were observed during the performance conditions compared to the perceptual condition, suggests that during the performance conditions, pianists did not need to wait until auditory feedback became available in order to judge the markings. Rather, they were able to produce accurate responses based on what they intended to play. During the performance conditions, verbal responses, on average, came midway through the note sequences that corresponded to dynamic and articulation markings rather than at the end of those sequences. Had knowledge of whether each marking matched their intentions not been verbalizable, or accessible to musicians' conscious awareness, until after they had performed the corresponding notes, verbal responses would have come several beats later than they did. Verbal task response times suggest that participants were anticipating their performance whether auditory and motor feedback were present or not.

A criticism levelled at many imagery studies is that the tasks used do not require depictive imagery to complete. Rather, responses may be generated on the basis of knowledge about the stimulus (Hubbard, 2010; Pylyshyn, 1981). In the context of the present experiment, this might involve a performer endorsing a diminuendo because they have identified a phrase ending and know that diminuendi are often found in that structural location. Had verbal responses derived from knowledge about musical structure in this way, however, faster average response times in the verbal compatibility judgement task than those observed would have been predicted, indicating a tendency for pianists to jump immediately to markings as they appeared without mentally representing the time course of the intervening music. Though responses were made faster during the performance conditions than during the perceptual condition, the average delay between the appearance of each marking and the production of a verbal response was approximately 5 to 7 beats for the Harthan piece and 7 to 9 beats for the Kabalevsky piece. These delays in response are in line with what might be expected if pianists were imagining the music at approximately the prescribed tempo. Furthermore, pianists were given only a short amount of time to prepare the pieces, and they did not know in advance which expressive markings they would be required to judge. This was done to avoid encouraging the encoding of expressive information as a list of declarative statements, which would not be an effective method for remembering expressive plans during normal performance, when the music tends to be more complex (see Chaffin et al., 2010; Chaffin & Logan, 2006). This experiment, therefore, provides evidence that imagery for dynamics and articulation can occur during physical performance as well as in the absence of sound and overt movement.

To our knowledge, these findings are the first indication that articulation, which has received little attention in the musical imagery literature, can be imagined. Perceived and produced articulation have previously been found to differ (Repp, 1995, 1998). The degree of detachment between notes produced by pianists instructed to play "optimally staccato," for example, is greater than the degree of detachment selected as 'optimally staccato' by listeners presented with the same passages of music (Repp, 1998). Such findings could indicate a shortcoming in pianists' ability to realize their intended articulation, perhaps due to an inability to specify articulation in plans, an inability to execute those plans, or ineffective monitoring of performance. The current study shows that articulation can be imagined during piano performance, and that these performance plans can be

executed whether or not auditory feedback is present. Like dynamics (Fabiani & Friberg, 2011), the perception of articulation is influenced by other factors, including global tempo, pitch register (Repp, 1998), and loudness (Hähnel & Berndt, 2010). Future research may reveal whether discrepancies between produced and perceived articulation derive from differences in how these factors contribute to planning and perception.

The fastest verbal task responses were observed during the motor-only performance, providing evidence that imagery for dynamics and articulation was stronger when auditory feedback was absent and motor feedback was present than during either the auditory-motor performance, when both types of feedback were present, or the imagined performance, when both were absent. It could be argued that verbal task performance was hampered in the auditory-motor condition by increased cognitive load resulting from the presence of and need to monitor auditory feedback; however, numbers of pitch and timing errors were comparable across auditory-motor and motor-only conditions. If the cognitive load for one condition was greater than for the other, differences in error rates would likely have been observed. Such an explanation would also imply that cognitive load is at a minimum during the imagined performance, when neither auditory nor motor feedback is available and needs to be monitored. The slower response times that were observed during the imagined performance compared to the motor-only performance, though, suggest that imagery may have been stronger when motor feedback was present. It could be that the removal of auditory feedback during performance encourages reliance on imagery. It may be more important to specify how actions are to be executed when less information is available about whether the desired effects have been achieved. An alternative explanation is that the presence of motor feedback facilitates auditory imagery during motor-only performance conditions. Research suggests that an auditory-motor integration network underlies musical imagery (Halpern, Zatorre, Bouffard, & Johnson, 2004; Hickok et al., 2003) and contributes to music performance (Keller et al., 2010; Keller & Koch, 2008). Whether auditory feedback, similarly, would facilitate motor imagery during disrupted motor feedback conditions is an interesting question for future research. Expressive parameters such as dynamics and articulation may be particularly tied to the motor modality as well. It has been proposed that the force or effort exerted by a performer is expressed through sound energy and largely perceived in terms of loudness (Dean & Bailes, 2008, 2010), and it may be more difficult to mentally represent loudness in the absence of motor activity.

#### THE RELATIONSHIP BETWEEN ONLINE IMAGERY ABILITY AND MUSICAL EXPERTISE

Within-subject consistency in expression across performances improved with increasing musical expertise. This relationship was significant for articulation and nearly significant for dynamics. Verbal task response times also improved, providing support for the predicted relationship between musical expertise and online imagery ability. It could be that experts' intentions regarding dynamics and articulation are more accessible to conscious awareness than novices'. Alternatively, previous research has shown that the range of planning, or the span of notes that are accessible at a given time, increases with musical expertise (Drake & Palmer, 2000; Palmer & Drake, 1997; Palmer & Pfordresher, 2003). The results of the current experiment suggest that expert musicians may have conscious awareness of a larger span of information relating to dynamics and articulation as well. The relationship between expertise and the consistency of performed dynamics and articulation is in line with previous research suggesting that expressive performance becomes more reliable with increasing expertise (Takahashi & Tsuzaki, 2008). A question deserving further research is whether expert imagery abilities are independent of expert technical abilities and therefore transferable across instruments. Four of the additional pianists who completed this experiment but were excluded from the analyzes met the criteria for inclusion in the novice pianist group, but had extensive training on another instrument as well. While the consistency of their dynamic and articulation profiles across conditions was similar to that of other novice pianists, their mean verbal response times were comparable to those of the experts.

Another question addressed by the present experiment is whether expert music abilities are music-specific or relate to general cognitive abilities. A potential relationship between imagery ability and working memory capacity was considered, as working memory mediates mental imagery (Aleman & Wout, 2004; Baddeley & Andrade, 2000) and has been predicted to co-vary with musical expertise (Franklin et al., 2008; Jakobson et al., 2003). No evidence was found for a relationship between working memory capacity and expertise in the current experiment, however; nor was working memory capacity found to explain imagery task performance. The effects of musical expertise observed in other studies of working memory may relate to their use of verbal tasks to assess working memory and the more effective rehearsal techniques that musicians can employ to aid their performance on such tasks (Franklin et al., 2008).

Verbal rehearsal is not likely to be a practical technique for use in an online imagery task, as the subvocalization that contributes to verbal rehearsal (Franklin et al., 2008) and the motor demands of music performance are likely to interfere with each other. In future research, the potential relationship between musical imagery ability and nonverbal working memory capacity should be investigated, as the ability to retain nonverbal auditory material (e.g., instrumental music; Bertz, 1995; Salamé & Baddeley, 1989), visual material (e.g., musical notation; Schendel & Palmer, 2007), or motor material (e.g., expressive gestures; Bailes, Bishop, Stevens & Dean, 2012) might contribute to success at imagining music.

#### THE ROLE OF SCHEMATIC EXPECTATIONS IN EXPRESSIVE PERFORMANCE PLANNING

Consciously accessible, anticipatory online imagery was assessed with the verbal compatibility judgement task in this experiment, but the results also emphasize the importance of unconscious expectations arising from familiarity with Western music structure. Pianists were significantly more likely to erroneously endorse markings taken from the original scores than to erroneously endorse markings designed to contradict the original scores. Several participants described being unsure at times of whether they were agreeing with a marking because it indicated something they had done during their baseline performance or because it was something that made sense in the context and something they might have done had they thought of it. The tendency to err in favor of original markings was observed in all expertise groups, though particularly apparent for experts during the motor-only and imagined conditions. For the pieces used in this experiment, therefore, the tendency to develop schematic expectations seems to have been shared by expert, intermediate, and novice pianists. The effect may be heightened for experts under conditions encouraging increased reliance on imagery. This finding accords with previous research suggesting that people with little or no music training may have a great deal of knowledge about musical structure, even if they lack the skills to show it on many musical tasks (Bigand & Poulin-Charronnat, 2006).

The process of interpreting a piece of music as it is prepared for performance has been suggested to be largely intuitive, with decisions about how to manipulate expressive parameters made in response to structural cues without the conscious awareness of the performer (Bangert et al., 2009). The contribution of unconscious schematic expectations to online planning does not preclude the use of consciously accessible imagery; the information represented in an image may

derive from either consciously controlled decisions about how a piece should go or unconscious expectations stimulated by structural cues. This experiment addressed the question of how expressive plans are represented rather than how they are created, and participants' abilities to verbally judge dynamic and articulation markings in advance of performing the corresponding notes suggest that expressive plans can be represented in the form of a consciously accessible image.

#### CONCLUSIONS

Expert musicians report relying on imagery to guide their expressive performance (Holmes, 2005; Rosenberg & Trusheim, 1990; Trusheim, 1993); however, little is known about how well parameters of expression can be imagined. The question of how online imagery contributes to the successful realization of expressive plans has likewise received little attention in the literature. The possibility that enhanced online imagery ability helps to explain experts' extraordinary control over expression was investigated in the current experiment with an assessment of online imagery abilities in novice, intermediate, and expert pianists. The results suggest that both dynamics and articulation can be imagined during performance: in all performance conditions, pianists made accurate verbal judgements of dynamic and articulation markings faster than they would have if they needed to wait until auditory feedback was available for retrospective evaluation. The results also suggest that the strength of this imagery improves with increasing musical expertise. Baseline dynamics and articulation were replicated with similar precision under auditory-motor and motor-only conditions, suggesting that in the absence of auditory feedback, participants could imagine the effects of their actions and compensate for the lack of sound. Aural skills are practiced by many music students, but if consciously imagining a desired sound increases the success with which intentions are realized, then an increased focus on developing and diversifying auditory imagery skills may be beneficial.

#### Author Note

This research was completed while Laura Bishop was a doctoral student and Freya Bailes was a Senior Lecturer 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.
- ALEMAN, A., & WOUT, M. (2004). Subvocalization in auditory-verbal imagery: Just a form of motor imagery? *Cognitive Processing*, 5, 228-231.
- BADDELEY, A. D., & ANDRADE, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General*, 129, 126-145.
- BAILES, F., & BISHOP, L. (2012). Musical imagery in the creative process. In D. Collins (Ed.), *The Act of Musical Composition* (pp. 55-77). Surrey, UK: Ashgate SEMPRES Psychology of Music Series.
- BAILES, F., BISHOP, L., STEVENS, C., & DEAN, R. T. (2012). Mental imagery for musical changes in loudness. *Frontiers in Perception Science*, 3. doi: 10.3389/fpsyg.2012.00525
- 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, Auckland, NZ.
- BEILock, S., BERTENTHAL, B. I., HOERGER, M., & CARR, T. H. (2008). When does haste make waste? Speed-accuracy tradeoff, skill level, and the tools of the trade. *Journal of Experimental Psychology: Applied*, 14, 340-352.
- BEILock, S., WIERENGA, S., & CARR, T. (2002). Expertise, attention, and memory in sensorimotor skill execution: Impact of novel task constraints on dual-task performance and episodic memory. *Quarterly Journal of Experimental Psychology Section A*, 55, 1211-1240.
- BERTZ, W. L. (1995). Working memory in music: A theoretical model. *Music Perception*, 12, 353-364.
- 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.
- BISHOP, L., BAILES, F., & DEAN, R. T. (2013). Musical expertise and the ability to imagine loudness. *PLOS ONE*, 8(2), e56052.
- BROWN, R. M., & PALMER, C. (2012). Auditory-motor learning influences auditory memory for music. *Memory and Cognition*, 40, 567-578.
- CHAFFIN, R., LISBOA, T., LOGAN, T., & BEGOSH, K. T. (2010). Preparing for memorized cello performance: The role of performance cues. *Psychology of Music*, 38, 3-30.
- CHAFFIN, R., & LOGAN, T. (2006). Practicing perfection: How concert soloists prepare for performance. *Advances in Cognitive Psychology*, 2, 113-130.
- CLARKE, E. (1993). Imitating and evaluating real and transformed musical performances. *Music Perception*, 10, 317-341.
- CONNOLLY, C., & WILLIAMON, A. (2004). Mental skills training. In A. Williamon (Ed.), *Musical excellence: Strategies and techniques to enhance performance*. (pp. 221-245). Oxford, UK: Oxford University.
- CONWAY, A., KANE, M., BUNTING, M., HAMBRICK, D., WILHELM, O., & ENGLE, R. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review*, 12, 769-786.
- 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.
- COWELL, H. (1926). The process of musical creation. *American Journal of Psychology*, 37, 233-236.
- DEAN, R. T., & BAILES, F. (2008). Is there a 'rise-fall temporal archetype' of intensity in electroacoustic music? *Canadian Acoustics - Acoustique Canadienne*, 36, 112-113.
- DEAN, R. T., & BAILES, F. (2010). A rise-fall temporal asymmetry of intensity in composed and improvised electroacoustic music. *Organised Sound*, 15, 147-158.
- DICKEY, M. R. (1991). A comparison of verbal instruction and nonverbal teacher-student modeling in instrumental ensembles. *Journal of Research in Music Education*, 39, 132-142.
- DRAKE, C., & PALMER, C. (2000). Skill acquisition in music performance: Relations between planning and temporal control. *Cognition*, 74, 1-32.
- DUKE, R. A., CASH, C. D., & ALLEN, S. E. (2011). Focus of attention affects performance of motor skills in music. *Journal of Research in Music Education*, 59, 44-55.
- FABIANI, M., & FRIBERG, A. (2011). Influence of pitch, loudness, and timbre on the perception of instrument dynamics. *Journal of the Acoustical Society of America*, 130, 193-199.
- FINNEY, S., & PALMER, C. (2003). Auditory feedback and memory for music performance: sound evidence for an encoding effect. *Memory and Cognition*, 31, 51-64.
- FRANKLIN, M., MOORE, S., YIB, C., JONIDES, J., RATTRAY, K., & MOHER, J. (2008). The effects of musical training on verbal memory. *Psychology of Music*, 36, 353-365.
- FRIBERG, A., & BATTEL, G. (2002). Structural communication. In R. Parncutt & G. McPherson (Eds.), *The science and psychology of music performance: Creative strategies for teaching and learning*. (pp. 199-218). New York: Oxford University.
- FURUYA, S., & SOECHTING, J. F. (2010). Role of auditory feedback in the control of successive keystrokes during piano playing. *Experimental Brain Research*, 204, 223-237.
- GIORGINO, T. (2009). Computing and visualizing dynamic time warping alignments in R: The dtw package. *Journal of Statistical Software*, 31(7), 1-24.

- HÄHNEL, T., & BERNDT, A. (2010, June). *Expressive articulation for synthetic music performances*. Paper presented at the Proceedings of New Interfaces for Musical Expression Conference, Sydney, Australia.
- HALPERN, A., ZATORRE, R., BOUFFARD, M., & JOHNSON, J. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia*, 42, 1281-1292.
- 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 memorised piano performance. *Bulletin of the Council for Research in Music Education*, 156, 58-65.
- HOLMES, P. (2005). Imagination in practice: A study of the integrated roles of interpretation, imagery and technique in the learning and memorisation processes of two experienced solo performers. *British Journal of Music Education*, 22, 217-235.
- HUBBARD, T. (2010). Auditory imagery: Empirical findings. *Psychological Bulletin*, 136, 302-329.
- JAKOBSON, L. S., CUDDY, L. L., & KILGOUR, A. R. (2003). Time tagging: A key to musicians' superior memory. *Music Perception*, 20, 307-313.
- 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.
- JUSLIN, P. (2003). Five facets of musical expression: A psychologist's perspective on music performance. *Psychology of Music*, 31, 273-302.
- JUSLIN, P., & LAUKKA, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, 129, 770-814.
- KALAKOSKI, V. (2001). Musical imagery and working memory. In R. I. Gødoy & H. Jørgensen (Eds.), *Musical imagery* (pp. 43-56). Lisse, The Netherlands: Swets & Zeitlinger.
- KEELE, S. W. (1968). Movement control in skilled motor performance. *Psychological Bulletin*, 70, 387-403.
- KELLER, P. (2012). Mental imagery in music performance: Underlying mechanisms and potential benefits. *Annals of the New York Academy of Sciences*, 1252, 206-213.
- KELLER, P., 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.
- KELLER, P., & KOCH, I. (2008). Action planning in sequential skills: Relations to music performance. *Quarterly Journal of Experimental Psychology*, 61, 275-291.
- KENDALL, R., & CARTERETTE, E. (1990). The communication of musical expression. *Music Perception*, 8, 129-163.
- KOCH, I., KELLER, P., & PRINZ, W. (2004). The ideomotor approach to action control: Implications for skilled performance. *International Journal of Sport and Exercise Psychology*, 2, 362-375.
- KUNDE, W., KOCH, I., & HOFFMAN, J. (2004). Anticipated action effects affect the selection, initiation, and execution of actions. *Quarterly Journal of Experimental Psychology*, 57A, 87-106.
- LASHLEY, K. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112-136). New York: Wiley.
- LAUKKA, P. (2004). Instrumental music teachers views on expressivity: A report from music conservatoires. *Music Education Research*, 6, 45-56.
- LINDSTRÖM, E., JUSLIN, P., BRESIN, R., & WILLIAMON, A. (2003). "Expressivity comes from within your soul": A questionnaire study of music students' perspectives on expressivity. *Research Studies in Music Education*, 20, 23-47.
- NAKAMURA, T. (1987). The communication of dynamics between musicians and listeners through musical performance. *Perception and Psychophysics*, 41, 525-533.
- OLLEN, J. (2006). *A criterion-related validity test of selected indicators of musical sophistication using expert ratings*. Unpublished doctoral dissertation, Ohio State University.
- PALMER, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 331-346.
- PALMER, C. (1997). Music performance. *Annual Review of Psychology*, 48, 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., & PFORDRESHER, P. (2003). Incremental planning in sequence production. *Psychological Review*, 110, 683-711.
- PALMER, C., & VAN DE SANDE, C. (1993). Units of knowledge in music performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 457-470.
- 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.
- 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., & MANTELL, J. T. (2012). Effects of altered auditory feedback across effector systems: Production of melodies by keyboard and singing. *Acta Psychologica*, 139, 166-177.
- PITT, M., & CROWDER, R. (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. (1995). Acoustics, perception, and production of legato articulation on a digital piano. *Journal of the Acoustical Society of America*, 97, 1878-1890.
- REPP, B. (1997). Expressive timing in a Debussy prelude: A comparison of student and expert pianists. *Musicae Scientiae*, 1, 257-268.
- REPP, B. (1998). *Perception and production of staccato articulation on the piano*. Retrieved from <http://www.haskins.yale.edu/haskins/STAFF/repp.html>
- REPP, B. (1999). Effects of auditory feedback deprivation on expressive piano performance. *Music Perception*, 16, 409-438.
- REPP, B. (2001). Expressive timing in the mind's ear. In R. Godøy & H. Jørgensen (Eds.), *Musical imagery* (pp. 185-200). Lisse, The Netherlands: Swets & Zeitlinger.
- ROSENBERG, H. S., & TRUSHEIM, W. (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.
- RUIZ, M., JABUSCH, H., & ALTENMÜLLER, E. (2009). Detecting wrong notes in advance: Neuronal correlates of error monitoring in pianists. *Cerebral Cortex*, 19, 2625-2639.
- SALAMÉ, P., & BADDELEY, A. (1989). Effects of background music on phonological short-term memory. *Quarterly Journal of Experimental Psychology*, 41, 107-122.
- SCHENDEL, Z. A., & PALMER, C. (2007). Suppression effects on musical and verbal memory. *Memory and Cognition*, 35, 640-650.
- SCHMIDT, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, 82, 225-260.
- SCHNEIDER, W., & SHIFFRIN, R. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1-66.
- STOCK, A., & STOCK, C. (2004). A short history of ideo-motor action. *Psychological Research*, 68(2-3), 176-188.
- 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.
- TRUSHEIM, W. (1993). Audiation and mental imagery: Implications for artistic performance. *Quarterly Journal of Music Teaching and Learning*, 2, 139-147.
- TURNER, M. L., & ENGLE, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28, 127-154.
- WILSON, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review*, 9, 625-644.
- 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.
- 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.
- WOODY, R. (2003). Explaining expressive performance: Component cognitive skills in an aural modeling task. *Journal of Research in Music Education*, 51, 51-63.