Memory and Metacognition for Piano Melodies: Illusory Advantages of Fixed over Random-Order Practice

Branden Abushanab and Anthony J. Bishara

College of Charleston

Author Note

Branden Abushanab is now at the Department of Psychology, University of North Carolina at Greensboro.

This project was completed as part of the requirements for the first-author's Bachelor's Essay at the College of Charleston. This project was supported by a College of Charleston SURF grant. We thank Major Murphy for developing the software used in this project.

Correspondence concerning this article should be addressed to Anthony J. Bishara, Dept. of Psychology, College of Charleston, 66 George St., Charleston, SC 29424. Phone: 843-953-5517. Fax: 843-953-7151. E-mail: BisharaA@cofc.edu

This paper is now in press:

Abushanab, B., & Bishara, A.J. (in press). Memory and metacognition for piano melodies: Illusory advantages of fixed over random-order practice. *Memory & Cognition*.

Abstract

Some learning schedules can foster an illusion of competence, whereby the learner feels that the skill will be better retained than it actually will be. Consider fixed-order practice, where a person practices a task repeatedly before switching to the next task (e.g., Task A, A, B, B); in contrast, in random-order practice, a person randomly alternates among two or more tasks (C, D, D, C). In the present experiment, participants (*n* = 25) who had formal training in piano practiced melodies under fixed or random-order conditions (within-subjects), and then returned for a retention test two days later. Initially, participants performed faster on melodies practiced in fixed-order. However, on a retention test two days later, participants were faster with melodies from the random-order condition. Despite the within-subjects design, which facilitates the comparison of practice conditions, participants' metacognitive judgments indicated an illusion of competence, whereby participants erroneously believed that fixed-order practice would result in faster retention performance. Our results suggest that even some trained musicians may use ease of acquisition as a heuristic for predicting future performance.

KEYWORDS: contextual-interference, metacognition, JOL, motor learning

Memory and Metacognition for Piano Melodies: Illusory Advantages of Fixed over Random-Order Practice

The order in which someone practices a motor skill, such as playing an instrument or typing, can have a large impact on the memory retention for that skill. Past research has revealed a paradox, which is that the practice schedule that produces superior performance at the time of acquisition may yield inferior performance at retention (Simon & Bjork, 2001; 2002). Over the years this paradox has been demonstrated in motor learning for various skills such as keyed timing (Simon & Bjork, 2001; 2002), golfing (Porter, Landin, Hebert, & Baum, 2007), knot tying (Ollis, Button, & Fairweather, 2005), and musical instrument learning (Stambaugh, 2011). Additionally, practice schedules can influence an individual's judgments of learning by fostering an *illusion of competence*, whereby the learner feels that the skill will be better retained than it actually will be. How robust is this illusion? The present report seeks to demonstrate both a) that the illusion can occur even in formally trained individuals, in this case, musicians, and b) that the illusion can occur even when participants can easily compare practice schedules.

When describing the influence of practice order on acquisition and retention of a motor activity, one can draw upon the contextual-interference principle (Battig, 1979; see Magill & Hall, 1990, for a review). The contextual-interference principle proposes that arranging to-be-learned material in a way that introduces interference between those materials typically impedes performance at acquisition, but frequently leads to superior performance at the time of retention (Lee & Simon, 2004). For example, imagine that an individual is asked to practice melodies A and B in a fixed-order (e.g., Melody A, A, A, B, B, B), and is asked to practice melodies C and D in a random-order (e.g., Melody C, D, C, C, D, D). Given that practicing melodies C and D requires the learner to alternate between both melodies in no specific pattern, arranging items in random-order is one way to introduce interference during acquisition of a skill (Shea & Morgan, 1979). Because of the high interference during learning of melodies C and D, the contextual-interference principle would predict poor acquisition and superior retention performances for those particular melodies. Additionally, the contextual-interference principle would suggest that judgments made while under the influence of contextual-interference may reflect misconceptions of future performance.

Research has revealed that judgments of learning (JOLs) are sometimes an unreliable heuristic for future performance (e.g., Bjork, 1999; Jacoby, Bjork, & Kelley, 1994; Koriat & Bjork, 2005). For example, JOLs are generally sensitive to manipulations that occur throughout learning, and certain manipulations may beget an illusion of competence, whereby learners feel that the skill will be better retained than it actually will be (Schmidt & Bjork, 1992; Simon & Bjork, 2001; 2002). To this end, JOLs made after a delay are often far more accurate at predicting subsequent performance than JOLs made immediately after study (Koriat & Bjork, 2005). One possible explanation of this pattern is that JOLs made immediately after study are made in the presence of the target (i.e., target is present in working memory), and predictions made at this time often reflect overconfidence (Koriat & Bjork, 2005). Therefore, given the notion that certain study conditions may receive more support from working memory (e.g. fixed-order practice) than others (random-order practice) (Lee & Magill, 1985), it can be presumed that the type of practice condition may influence a learner's prediction of future performance.

JOLs can be measured in various ways, including asking participants which practice schedule they would prefer to use in the future (Kornell, Castel, Eich, & Bjork, 2010) or requiring participants to judge how fast they believe they can later execute a task (Simon & Bjork 2001; 2002). When using preference JOLs as an assessment of learning, individuals significantly prefer fixed over random-order practice regardless of retention outcome (Kornell et al., 2010). This misconception in learning is also reflected in time-based JOLs, whereby participants exhibit an illusion of competence for tasks practiced in fixed-order (Simon & Bjork, 2001; 2002). Such JOL miscalibration may be due to the use of retrieval fluency as a basis for predicting future performance (Koriat & Ma'ayan, 2005). In other words, the initial ease of performance of fixed-order practice may lead to overconfidence, and it may do so on a variety of metacognitive judgments.

The interaction between metacognition and type of study has been closely examined for motor learning. In a particularly relevant set of experiments (Simon & Bjork, 2001; 2002), participants learned keystroke patterns on the number pad of a computer by either fixed-order practice trials (e.g., pattern A, A, A, B, B, B, C, C, C) or random-order practice trials (e.g., pattern D, E, D, E, F, D, F, E, F). Each to-be-learned pattern had a specific goal movement time, and participants were instructed to execute each pattern within its specified goal movement time. After the acquisition phase, participants were asked to make predictions of how close they thought their performance would be to the relevant goal movement time of each pattern when given a retention test the next day. Results from the acquisition phase showed that participants who learned patterns under fixed-order practice outperformed those who learned patterns under random-order practice. However, results from the retention phase showed the opposite pattern: those who had learned with random-order practice performed superior to those who had learned with fixed-order practice. In regards to participants' JOLs, those who learned under fixed-order practice showed an illusion of competence in that they believed their future performances would be better than they actually were. Thus, the type of practice schedule used during acquisition had a sizeable effect on future retention performance, and JOLs were poor predictors of later performance.

Contextual-interference effects have been observed even with real-world scenarios such as with batting by baseball players (Hall, Domingues, & Cavazos, 1994) and knot-tying by fire-fighters (Ollis, Button, & Fairweather, 2005). One previous study has manipulated fixed versus random-order practice with novice musicians (Stambaugh, 2011). Replicating the typical behavioral pattern from the contextual-interference literature, it was found that fixed-order practice led to superior acquisition but inferior retention of clarinet pitch exercises. However, the influence of contextual-interference on musicians' metacognition has yet to be examined. Revealing that musicians may be susceptible to the illusion of competence produced by fixed-order practice could have serious implications for music pedagogy. For example, misconceptions at learning may alter a musician's control over their own, self-regulated practice by leading them to spuriously endorse an inferior practice routine (fixed-order).

Perhaps more importantly, previous research has shown an illusion of competence in between-subjects designs (Simon & Bjork, 2001; 2002), but such designs might not be the most challenging ones for illusions. Different cues are made accessible in between-subject and in within-subject experimental designs (Kahneman & Tversky, 1996). In particular, within-subject designs measure joint evaluations, thereby making the differences among conditions more salient (Kahneman, 2002). For this reason, cognitive illusions sometimes disappear when between-subject tasks are extended to within-subject designs (Kahneman, 2002; Kahneman & Tversky, 1996; LeBoeuf & Shafir, 2003). In other words, within-subject designs may provide more stringent tests of the robustness of an illusion.

To the best of our knowledge, the illusion of competence observed in contextual-interference experiments has yet to be extended to a within-subjects design. On the one hand, a within-subjects design should facilitate participants' ability to directly compare fixed and random-order practice, perhaps eliminating the illusion of competence. However, such a direct comparison may provide little help if the cues used during this comparison are systematically biased against an accurate prediction of retention. If participants are using ease of acquisition performance as a heuristic for predicting retention, this heuristic would still suggest to the participant that fixed is better than random-order practice for later retention. That is, even when able to directly compare the conditions, acquisition performance may continue to thwart cognizant predictions of future performance.

Finding that the illusion of competence remains under the stringent test of a within-subjects setting would not only reveal the robustness of the illusion, but it would also have implications for music pedagogy. Case studies on musicians' self-regulated practice have revealed that, when left to their own devices, musicians often experiment with different practice schedules (Chaffin & Logan, 2006; Hallam, 2001; Nielsen, 1999; Noice, Jeffrey, Noice, & Chaffin, 2008). Therefore, testing the influence of fixed and random-order practice on metacognitive judgments in a within-subjects setting would be more analogous to real-world practice routines often adopted by musicians.

In the present report, participants practiced melodies under both fixed- and random-order practice conditions. Throughout this acquisition phase, they made predictions about their future performances after a 2-day delay. After the 2-day delay, participants returned and made predictions again about their expected performances, and then completed a final retention test in which they performed previously studied melodies under both fixed and random-order trial presentations. Additionally, at the end of each day, participants were asked to report their preference for fixed or random-order practice on a Likert-type scale. We hypothesized that, of the two practice types, fixed-order practice would lead to quicker performances during acquisition, but slower performance after the 2-day delay. Additionally, it was hypothesized that participants would incorrectly predict that fixed-order practice would lead to better (faster) retention, and would likewise prefer fixed-order practice.

Method

Participants

Participants were 25 individuals (6=Male, 19=Female) recruited from both the campus and the surrounding community of Charleston, SC. Recruitment techniques included newspaper advertisements (both print and electronic), classroom announcements, and fliers. Participants were 18 to 49 years of age (*M* = 25.8, *SD* = 10.1). All had 1 to 17 years of formal piano training (*M* = 8.0, *SD* = 3.9). Most participants (80%) had formal training on an additional instrument or voice (*M* = 2.6 years, *SD* = 2.7). Total formal music training (all instruments combined) ranged from 3.5 to 18 years (*M* = 10.6 years, *SD* = 3.9). Of those who had training on one or more additional instruments, 35% listed wind instruments, 25% voice, 15% guitar, 15% viola/violin, 15% organ, and 5% percussion. 10% of those participants reported having training on two additional instruments. 24% of participants reported being professional musicians, but professionals did not have significantly more years of training than non-professionals for either piano or all instruments combined, *p*s > .10.

All testing occurred within 2 sessions separated by a 2-day delay, with the first session lasting one hour and the second session lasting 30 minutes. Participants were tested individually in a private room with an experimenter and were compensated for their time at a rate of $10 per session.

Materials & Design

The experimental procedures were administered using a Yamaha P-200 digital piano connected via MIDI to a desktop computer. A customized software program was used to administer the experimental trials and collect participants' responses for later analysis. All participants were given headphones to wear with volume that they could adjust to a comfortable level. 12 melodies were taken from an obscure source (Ottman, 1996) and edited such that each note had the same duration (i.e. quarter notes) and could be executed for ease of right-hand performance. The melodies' key signatures had no more than 2 sharps or 2 flats. Each to-be-learned melody was 12 notes in duration and each had a unique goal movement time that was set in such a way so as to motivate fast performance, but participants were encouraged to surpass a melody's goal movement time if possible.

A within subjects design was adopted, with 4 melodies practiced in fixed-order (e.g., Melody A, A, A, B, B, B), and 4 practiced in random-order (e.g., Melody C, D, C, D, D, C). There were 8 critical melodies used throughout the acquisition phase (4 under fixed-order, 4 under random-order) and 1 melody used for the practice phase (before the acquisition phase). Additionally, there were 2 buffer melodies presented at the beginning and 1 presented at the end of the acquisition phase (Day 1) to prevent or reduce primacy and recency effects. The same 8 critical melodies used for the acquisition phase were used in the retention phase (48-hours later). The retention phase began with the practice melody and 2 buffer melodies, and concluded with 2 buffer melodies. Presentation of critical melodies was divided into 4 blocks, and each block contained 2 melodies presented in either fixed- or random-order. The presentation order of melodies was fixed, and the assignment of melodies to study condition was counterbalanced across participants.

The main dependent measure was participants' performance time on trials where all notes were played in the right order. Trials during which an incorrect note was played were rare (*M* = 6.3% of critical trials) and were excluded from analyses. Participants' time-based JOLs were collected after every fifth trial of each of the melodies, whereas participants' preference- and belief-based JOLs were collected at the ending of both days.

In addition to manipulating fixed versus random-order practice during acquisition, practice order was also manipulated during retention, with half of the melodies assigned to fixed and half assigned to random-order during retention. Assignment of melodies to retention conditions was also counterbalanced across participants. This retention manipulation was included so as to consider the possibility that the effect of acquisition practice order depended on retention practice order (i.e., the possibility that practice order benefits would be transfer-appropriate). However, consistent with previous research (Simon & Bjork, 2001; 2002), practice order during retention did not qualitatively alter the pattern of results (i.e., retention practice order did not significantly interact with acquisition practice order, $F<1$, $p = .48$). All other analyses collapsed across this factor.

Procedure

Acquisition Phase. At the start of the acquisition phase, participants were verbally instructed to play melodies as quickly as possible while maintaining 100% accuracy. Before each trial, the goal movement time for the particular melody (e.g., 3.4 seconds) appeared on the computer screen above the sheet music of the to-be-learned melody. The sheet music and goal movement time remained on the screen until the participant completed that trial according to the relative practice schedule. Additionally, before each of the trials, participants were reminded that they should prioritize accuracy, completing the melody as accurately as possible even if required to play at a slower speed in the beginning. Participants were also reminded that once accuracy was stable, they should attempt to increase their speed to get closer to, or even surpass, the goal movement time. After each trial, feedback was displayed that provided three pieces of information: whether or not the keys were pressed correctly (i.e. correct notes were displayed in green and incorrect notes were displayed in red), the goal movement time, and the time the participant took. If a melody was executed with 100% accuracy, the feedback screen instructed participants to try to increase their performance speed.

After 5 executions of each melody, participants were asked to make a prediction about how fast they believed they could execute that melody 2-days later, assuming they had no more practice in the meantime. At the end of the acquisition phase, participants were asked: *if you had to learn new melodies in the future which practice schedule would you prefer to use: fixed or random-order practice*. Participants responded to this question using a Likert-type scale (1= *strongly prefer mixed*, 2= *somewhat prefer mixed*, 3= *no preference*, 4= *somewhat prefer unmixed,* and 5= *strongly prefer unmixed*). Additionally, participants were provided with examples of both practice orders for clarification. Once completed, participants were instructed to return after 2 days at the same time as their original arrival to complete the retention phase of the experiment and were advised not to practice any of the melody patterns they were exposed to.

Retention Phase (2 Days Later). Upon returning, participants were handed a packet that contained screenshots of each of the melodies' corresponding sheet music and goal movement time. Participants were then asked for each melody: "How quickly do you think you will be able accurately execute this melody within 6 trials?" Following this, participants completed 6 trials of each of the previously studied melodies. On these retention test trials, consistent with the acquisition phase, the goal movement time was displayed above the sheet music of the to-be-performed melody. After each trial, before the next to-be-performed melody appeared, participants received the same feedback screen, which provided the same information as the acquisition phase. At the end of the retention phase, participants were asked three questions: 1) "if you had to learn melodies in the future which practice schedule would you prefer to use: fixed or random-order practice?", 2) "which melodies do you believe you performed best on today, those that you practiced on day 1 in a random order or those that you practiced on day 1 in fixed order?", and 3) "why do you believe this?" The first two questions used a 5-point Likert-type scale (see Acquisition Phase). Question 3 was open-ended and intended for exploratory purposes. Questions 2 and 3 were added after data collection began, and so results were missing for 3 participants for these questions.

Results

When sphericity was significantly violated - as shown by Mauchly's test - the Greenhouse-Geisser corrected statistics are reported.

Performance Time

Acquisition Phase. Acquisition performance times were subjected to a 2 × 6 (Practice Type × Trial Block) within-subjects analysis of variance (ANOVA). Melodies practiced in fixed-order were performed significantly faster than melodies practiced in random-order (see Figure 1A), as revealed by a significant main effect of practice type, $F(1,24)=10.44$, $p < .01$, $\eta_p^2 = .30$. Melodies were performed significantly faster as participants progressed through the trial blocks, as indicated by a significant main effect of trial block, $F(1.86,44.59)=132.93$, $p < .01$, η_p^2

= .85. Additionally, the Practice Type × Trial Block interaction was not significant, $F(2.64, 63.27)=1.03$, $p > .10$.

Retention Phase (2 Days Later). Melodies originally practiced in random-order were performed significantly faster than melodies originally practiced in fixed-order (see Figure 1A), as shown by a paired-samples t-test, $t(24) = 3.03$, $p < .01$. In order to compare retention phase to acquisition phase performance, participants' retention performances were then subjected to a 2 × 2 (Practice Type × Day) within-subjects ANOVA. Melodies practiced in fixed-order were significantly faster than melodies practiced in random-order on acquisition day (collapsing across all blocks); however, on the retention day melodies originally practiced in random-order outperformed the melodies practiced in fixed-order, as revealed by a significant crossover interaction effect, $F(1,24)=17.72$, $p < .01$, $\eta_p^2 = .43$. Moreover, that interaction was still significant even when retention day performance was compared to only the final block of acquisition day, $F(1,24)=17.75$, $p < .01$, $\eta_p^2 = .43$.

Metacognition

Acquisition Phase. As shown in Figure 1, participants' JOLs during acquisition revealed a pattern strikingly similar to participants' acquisition performance speeds. That is, predicted performance times for melodies practiced in fixed-order showed a similar pattern to participants' performance times on acquisition day in that those melodies were predicted to be performed faster after a 2-day delay, as revealed by a 2 × 6 (Practice Type × Trial Block) within-subjects ANOVA, $F(1,24)=6.72$, $p < .05$, $\eta_p^2 = .22$. As shown in Figure 1B, the predictions for future performance times for melodies practiced in fixed-order decreased as participants progressed through the trial blocks (similar to their performance), which was indicated by a significant main effect of practice type, $F(1.86,44.59)=132.93$, $p < .01$, $\eta_p^2 = .85$. The Practice Type × Trial Block interaction was not significant, $F(1.89, 14.03)=1.20$, $p > .10$.

Before Retention Phase. Participants' JOLs made at the beginning of the retention phase indicated no significant preference for practice type, as shown by a paired-samples t-test, $p > .50$. However, this finding may have been due to participants' inability to recall their previous performance times for each melody after the 2-day delay. A 2 × 2 (Practice Type × Day) within-subjects ANOVA revealed a significant interaction effect, $F(1,24)=6.11$, $p < .05$, $\eta_p^2 = .20$, showing that the pattern of JOLs for fixed versus random conditions changed from acquisition day to retention day. Moreover, that interaction remained significant even when comparing the JOLs for the final trial block of acquisition day to the beginning of the retention day JOLs, $F(1,24)=5.09$, $p < .05$, $\eta_p^2=.18$.

Preferences and Beliefs. Participants significantly preferred fixed-order practice, and there were non-significant trends in the same direction regarding participants' beliefs. Recall that participants responded on Likert scales ranging from 1 (strong preference/belief favoring random-order) to 5 (strong preference/belief favoring fixed-order). At the end of acquisition day, participants on average preferred fixed-order practice (M=3.7, SD=1.4). A one-sample t-test showed that the mean rating was significantly higher than the neutral rating of 3, $t(24)=2.37$, $p < .05$. Likewise, at the end of retention day, participants still significantly preferred fixed-order practice (M=3.6, SD=1.1), $t(24)=2.50$, $p < .05$. At the end of retention day, there was also a non-significant trend such that participants believed that fixed-order practice was superior (M=3.5, SD=1.2), $t(21)=1.74$, $p < .10$. In order to confirm these results without assuming an interval response scale, one-sample Wilcoxon signed rank tests were also performed, testing whether medians (rather than means) were significantly different from the neutral value of 3. As shown by signed rank tests, medians were significantly higher than 3 for preference judgments both at the end of the acquisition day, $p < .05$, and at the end of the retention day, $p < .05$. At the end of retention day, there was also a non-significant trend such that participants believed in the superiority of fixed-order practice, $p < .09$. The medians for all three questions were 4.0, consistent with a “slight” preference or belief in favor of fixed-order practice. Overall, participants’ preferences and beliefs were broadly consistent with their mistaken predictions made during the acquisition phase.

Effects of Formal Training

In order to explore the effects of formal training, all previously reported ANOVAs and t-tests were repeated with years of formal training in piano used as a covariate in ANCOVA. Whereas years of formal piano training was preserved as a continuous variable for inferential statistics, in order to present summary information, a median split was also conducted. The median years of formal piano training was 9 years, and so participants were divided into High Training Individuals (more than 9 years of training, n=10) and Low Training Individuals (less than 9 years, n=12). Results with High and Low Training groups are shown in Figure 2. Overall, while differences between fixed and random-order may have been somewhat reduced by training (often non-significantly reduced), there was no clear evidence that highly trained musicians showed a reversal of any patterns of interest. In particular, there was no clear evidence that highly trained individuals correctly predicted faster performance times for random-order practice.

Not surprisingly, more training was associated with shorter performance times (Figure 2A versus 2B), and likewise, JOLs that predicted shorter performance times on the retention test (Figure 2C versus 2D). In a 2 × 2 (Practice Type × Day) ANCOVA on performance time, years of formal training showed a significant effect, $F(1,23)=9.22$, $p < .01$, $\eta_p^2 = .29$. A 2 × 2 ANCOVA on JOLs also revealed a significant effect of year of formal training, $F(1,23)=4.85$, $p < .05$, $\eta_p^2 = .17$. Thus, more highly trained musicians played faster and predicted that they would play faster.

Interactions between Training and Performance Time. Of more importance, though, is whether training qualitatively altered the pattern of results, that is, whether training interacted with other variables. For acquisition performance times (Figure 2A and 2B), a 2 × 6 (Practice Type × Trial Block) ANCOVA showed a significant interaction between Training and Trial Block, suggesting that training affected the overall rate of learning during acquisition, $F(1.8, 41.0)=3.74$, $p < .05$, $\eta_p^2 = .14$. Examining performance times across the last block of acquisition day and the retention day, a 2 × 2 (Practice Type × Day) ANCOVA showed a significant interaction between Training and Day, consistent with more highly trained individuals forgetting less (i.e., slowing down less) between the acquisition and retention days, $F(1, 23)=4.33$, $p < .05$, $\eta_p^2 = .16$. All other interaction effects for performance time were not significant, *p*s > .05.

Interactions between Training and Predicted Retention Performance Time. In terms of time-based JOLs (Figure 2C and 2D), interaction effects with training were apparent only on the acquisition day. A 2 × 6 (Practice Type × Trial Block) ANCOVA on JOLs during the acquisition day showed a significant interaction between Training and Trial Block, $F(1.5, 34.2)=7.93$, $p < .01$, $\eta_p^2 = .26$. This significant interaction is analogous to one observed for Retention Performance, and it suggests that training affected the overall rate of learning, which in turn affected JOLs. Additionally, there was a three-way interaction between Training, Trial Block, and Practice Type, $F(2.2,50.5)=6.37$, $p < .01$, $\eta_p^2 = .22$. This three-way interaction is consistent with the pattern that low training individuals had larger JOL differences between fixed and random-order for earlier blocks than for later blocks, but such differences were very small for all blocks for high training individuals (see Figure 2C and D). All other interaction effects for JOLs were not significant, *p*s > .05.

Relationship of Training to Preferences and Beliefs. On Likert-scale responses about preferences and beliefs, there was some evidence that more highly trained individuals showed less preference for fixed-order practice, particularly at the end of the acquisition day. All Pearson correlations between training and preferences/beliefs were negative, indicating that training might sway participants away from fixed-order practice, but none of these correlations reached significance, *p*s > .05. Specifically, training was negatively correlated with the preference rating on acquisition day ($r = -.35$), retention day (-.14), and the belief rating on retention day (-.15). These correlations may have been underpowered due to the large number of tied data points, as Likert responses could only take on 5 possible values. To address this issue, correlations were also analyzed with Kendall's Tau B. These analyses showed a significant correlation between training and acquisition day preference ratings, $\tau_b = - .41$, $p < .05$, but not between training and retention day preferences, $\tau_b = -.15$, $p > .10$, or beliefs, $\tau_b = -.15$, $p > .10$. In other words, there was some evidence that training was associated with less preference for fixed-order practice, but only on the acquisition day. However, even among the High Training group with more than 9 years of formal training, mean Likert-scale responses never fell below 3.0 midpoint of the scale ($M = 3.0$, 3.2, and 3.1 for acquisition day preferences, retention day preferences, and retention day beliefs, respectively). Hence, training was associated with less preference for fixed-order practice, but the High Training group was, at best, agnostic about the benefits of random-order practice.

Discussion

Consistent with previous research (Stambaugh, 2011), we replicated the contextual-interference principle in the learning of a musical instrument; fixed-order practice fostered faster acquisition performance, whereas random-order practice resulted in faster retention performance. This finding suggests that random-order practice is more advantageous for long-term retention for piano melodies.

More importantly, participants showed the illusion of competence engendered by fixed-order practice in a within-subjects setting. To our knowledge, this is a novel finding. Participants' performance times during acquisition were aligned with their metacognitive judgments, suggesting that, even when they could easily compare fixed and random-order practice conditions, participants still erroneously espoused the short term benefits of fixed-order practice (cf. Simon & Bjork, 2001; 2002). This pattern was also in participants' preference-based judgments, which showed that, on both days, the majority of participants still favored fixed-order over random-order practice for future training.

Why did the illusion of competence prevail in a within-subjects setting? After all, within-subject designs often increase the use of controlled processes by allowing comparison among conditions (Kahneman, 2002). Additionally, our participants were provided with feedback, which should also increase the use of cognitive control (Kahneman & Klein, 2009). Increased cognitive control should allay some illusions by reducing the need to rely on heuristics, including retrieval fluency (Jones & Jacoby, 2001; also see Jacoby, Bishara, Hessels, & Hughes, 2007; Jacoby, Hessels, & Bopp, 2001). On the other hand, even though cognitive control may allow for easier comparison across conditions, in our scenario, participants may have been comparing conditions based on immediate performance speed. Facilitating this comparison would, if anything, contribute to illusions about retention, because immediate performance poorly predicted retention performance (cf. Koriat & Bjork, 2005).

In order to avoid the illusion of competence, a learner might need to compare retention performance, and particularly, retention performance as a function of acquisition type. However, this information may prove difficult for the learner to extract. The learner would not only need to notice retention performance differences, but would also need to remember how retention performance mapped on to acquisition type, sometimes from an acquisition that occurred several days earlier. To add to the conundrum, different stimuli (in this case, musical passages) often have different inherent levels of difficulty, thus making it even more challenging for learners to correctly attribute performance differences to practice order. Indeed, the benefits of random over fixed-order practice may be challenging to notice from casual, nonscientific experimentation with practice orders, and that may be the reason why even experts sometimes rely on fixed-order practice (Maynard, 2006).

Another challenge learners might experience may be that it seems counterintuitive to endorse the practice order associated with difficult acquisition and to ignore the ease of acquisition associated with fixed-order practice. This would suggest that participants may have been using acquisition performance as a heuristic for predicting both retention performance and preference-based JOLs. However, if participants were using retrieval fluency at learning to formulate their metacognitions, then why did participants still prefer fixed-order practice for future training at the end of the retention day, thereby disregarding the subpar retention performance for melodies practiced in fixed-order? There are at least two possible explanations for this pattern. First, participants could have been motivated to maintain consistency in their judgments (Abelson, Aronson, McGuire, Newcomb, Rosenberg, & Tannenbaum, 1968), and this may have led them to persist with their preferences made at the end of the first day. Second, after the 2-day delay, participants may have been unable to recall which melodies appeared in fixed or random-order practice conditions during acquisition, which would deter them from comparing acquisition practice schedules to their resulting retention performance speeds. The latter possibility would explain why participants' time-based JOLs made at the beginning of the retention day did not suggest a practice type preference. However, further research is required to disentangle these possibilities.

In the present experiment, metacognition was assessed at both the item-level (e.g., predictions of retention performance for each melody) and the global-level (e.g., preference for fixed or random-order practice in general). Item and global-level assessments do not always converge; even when item-level assessments poorly reflect actual performance, global-assessments are sometimes more accurate (e.g., Dunlosky & Hertzog, 2000; Koriat, Bjork, Sheffer, & Bar, 2004; Wahlheim, Finn, & Jacoby, 2012). In contrast, in the present experiment, both item-level and global-level assessments suggested poor awareness of the benefits of random-order practice. In particular, predicted retention performance times and more global preference ratings both favored fixed-order. The convergence among item-level and global-level measures underscores the robustness of the illusion of competence in trained musicians.

Of course, "trained" is a relative word, and some musicians had more training than others. Musicians with more training showed, if anything, smaller differences between fixed and random-order practice (see Figure 2), but such interaction effects often failed to reach significance. The sample may have been too small to examine the impact of a between-subject variable (years of formal piano training) with adequate power. Additionally, it is unclear if any such reduction in the difference between fixed and random-order practice would reflect better metacognition, or would simply reflect an approach toward asymptotic performance. That is, as performance times were smaller, differences between the two conditions were also smaller, both for high and low training individuals. Importantly, even individuals with more than 9 years of formal training were agnostic about the value of random-order practice. Overall, the pattern of results suggests that training might have reduced the illusion of competence, but it did not produce a clear preference for random over fixed-order practice.

One limitation of the present experiment is that the melodic stimuli were artificially simple. For example, our stimuli were comprised of only quarter notes, but real melodies often contain a variety of note durations and rhythms. Previous research with other motor skills suggests that contextual-interference's behavioral effects remain even with more complex stimuli (Ollis et al., 2005), but to the best of our knowledge, there is no empirical work investigating the interaction between contextual-interference and task complexity on metacognition.

Another limitation is the imbalance between conditions regarding elapsed time between item-level JOLs made during acquisition. Recall that the number of stimulus repetitions between JOL assessments was held constant, with JOL assessment occurring on every 5th presentation of a particular melody. However, equating melody repetitions between JOLs necessarily led to different amounts of time elapsing between JOLs for fixed versus random-order practice conditions. It should be noted that this confound between inter-JOL time and practice schedule is common in experiments of this nature (e.g., Simon & Bjork, 2001; 2002), perhaps because equating inter-JOL time would require confounding practice type with arguably more critical variables, such as the number of stimulus repetitions between JOLs. In any event, one possible concern with this procedure is that more time elapses between JOL assessments of the same stimulus in the random-order condition, which could cause participants to be less certain about their rate of learning across trials. While possible, this explanation of the current JOL results seems unlikely. In the fixed-order practice condition, JOL assessments occurred after 5 *consecutive* trials, which should, if anything, provide a more accurate gauge for rate of learning, yet this condition led to less accurate predictions for future performance. Additionally, participants were instructed to estimate their future performance assuming that they would have no additional practice, and so the rate of learning across trials should be a less relevant heuristic than performance on the most recent trial.

The findings from the current report have several implications. Most importantly, they reveal that the illusion of competence is robust enough to prevail in a within-subjects design. Second, they suggest that fixed-order practice can have deleterious effects on both the long-term retention and metacognition of trained musicians. Lastly, our results advise caution to musicians and other experts engaging in self-regulated practice. Even when such individuals casually experiment with both fixed and random-order practice, their metacognitive judgments may still be contaminated by the illusion of competence engendered by fixed-order practice, thus encouraging the use of inferior learning strategies.

References

Abelson, R. P., Aronson, E., McGuire, W. J., Newcomb, T. M., Rosenberg, M. J., & Tannenbaum, P. (Eds). (1968). *Theories of cognitive consistency: A sourcebook.* Chicago: Rand-McNally.

Battig, W. F. (1979). The flexibility of human memo In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory*. Hillsdale, NJ: Erlbaum. 23-44.

Bjork, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII. Cognitive regulation of performance: Interaction of theory and application* (pp. 435-459). Cambridge, MA: MIT Press.

Chaffin, R., & Logan, T. (2006). Practicing perfection: How concert soloists prepare for performance. *Advances in Cognitive Psychology*, *2*, 113–130.

Dunlosky, J., & Hertzog, C. (2000). Updating knowledge about encoding strategies: A componential analysis of learning about strategy effectiveness from task experience. *Psychology and Aging*, *15*(3), 462-474.

Hall, K. G., Domingues, D. A., & Cavazos, R. (1994). Contextual interference effects with skilled baseball players. *Perceptual and Motor Skills*, *78*, 835-841.

Hallam. S. (2001). The development of metacognition in musicians: Implications for education. *British Journal of Music Education*, *18*, 27-39.

Jacoby, L. L., Bishara, A. J., Hessels, S., & Hughes, A. (2007). Probabilistic retroactive interference: The role of accessibility bias in interference effects. *Journal of Experimental Psychology: General*, *136*, 200-216.

Jacoby, L. L., Bjork, R. A., & Kelley, C. M. (1994). Illusions of comprehension, competence, and remembering. In D. Druckman & R. A. Bjork (Eds.), *Learning, remembering,* *believing: Enhancing human performance* (pp. 57-80). Washington, DC: National Academy Press.

Jacoby, L. L., Hessels, S., & Bopp, K. (2001). Proactive and retroactive effects in memory performance: Dissociating recollection and accessibility bias. In H. Roediger, J. S. Nairne, I. Neath, A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 35-54). Washington, DC US: American Psychological Association.

Jones, T. C., & Jacoby, L. L. (2001). Feature and conjunction errors in recognition memory: Evidence for dual-process theory. *Journal of Memory and Language, 45*, 82-102.

Kahneman, D. (2002). Maps of Bounded Rationality: A Perspective on Intuitive Judgment and Choice, Nobel Prize Lecture, December 8, Stockholm

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise. *American Psychologist, 64,* 515–526.

Kahneman, D. & Tversky, A. (1996). On the reality of cognitive illusions: A reply to Gigerenzer's critique. *Psychological Review, 103,* 582-591.

Koriat, A., & Bjork, R.A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 187–194.

Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General, 133*(4), 643-656.

Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language, 52,* 478–492.

Kornell, N., Castel, A. D., Eich, T. S., & Bjork, R. A. (2010). Spacing as the friend of both memory and induction in young and older adults. *Psychology and Aging*, *25*, 498–503.

LeBoeuf, R. A., & Shafir, E. (2003). Deep thoughts and shallow frames: On the susceptibility to framing effects. *Journal of Behavioral Decision Making*, *16*, 77–92.

Lee, T. D., & Magill, R. A. (1985). Can forgetting facilitate skill acquisition? In D. Goodman, R. B. Wilberg, & I. M. Franks (Eds.), *Differing perspectives in motor learning, memory, and control*. Amsterdam: Elsevier. 3-22.

Lee, T. D., & Simon, D. A. (2004). Contextual interference. In A. M. Williams & N. J. Hodges (Eds.), *Skill acquisition in sport: Research, theory and practice.* London: Routledge. 29-44.

Magill, R.A., & Hall, K.G. (1990). A review of the contextual interference in motor-skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 730-746.

Maynard, L. (2006). The role of repetition in the practice sessions of artist-teachers and their students. *Bulletin of the Council for Research in Music Education*, *167*, 61–72.

Nielsen, S.G. (1999). Learning strategies in instrumental music practice. *British Journal of Music Education, 16,* 275-291.

Noice, H., Jeffrey, J., Noice, T., & Chaffin, R., (2008). Memorization by a jazz pianist: A case study. *Music Psychology*, *36*, 63–79.

Ollis S., Button C., & Fairweather, M. (2005). The influence of professional expertise and task complexity upon the potency of the contextual interference effect. *Acta Psychol (Amst), 118*, 229–44.

Ottman, R. W. (1996). *Music for sight singing* (4th ed.). Englewood Cliffs, NJ: Prentice-Hall

Porter, J. M., Landin, D., Hebert, E. P., & Baum, B. (2007). The effects of three levels of contextual interference on performance outcomes and movement patterns in golf skills. *International Journal of Sports Science & Coaching*, *2,* 243.

Schmidt, R.A., & Bjork, R.A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*, 207–17.

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 179-187.

Simon, D. A., & Bjork, R. A. (2001). Metacognition in motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 907-912.

Simon, D. A., & Bjork, R. A. (2002). Models of performance in learning multi-segment movement tasks: Consequences for acquisition, retention and judgments of learning. *Journal of Experimental Psychology: Applied*, *8*, 222-232.

Stambaugh, L. A. (2011). When repetition isn't the best practice strategy: Effects of blocked and random practice schedules. *Journal of Research in Music Education*, *58*, 368-383.

Wahlheim, C. N., Finn, B., & Jacoby, L. L. (2012). Metacognitive judgments of repetition and variability effects in natural concept learning: Evidence for variability neglect. *Memory & Cognition*, *40*(5), 703-716.

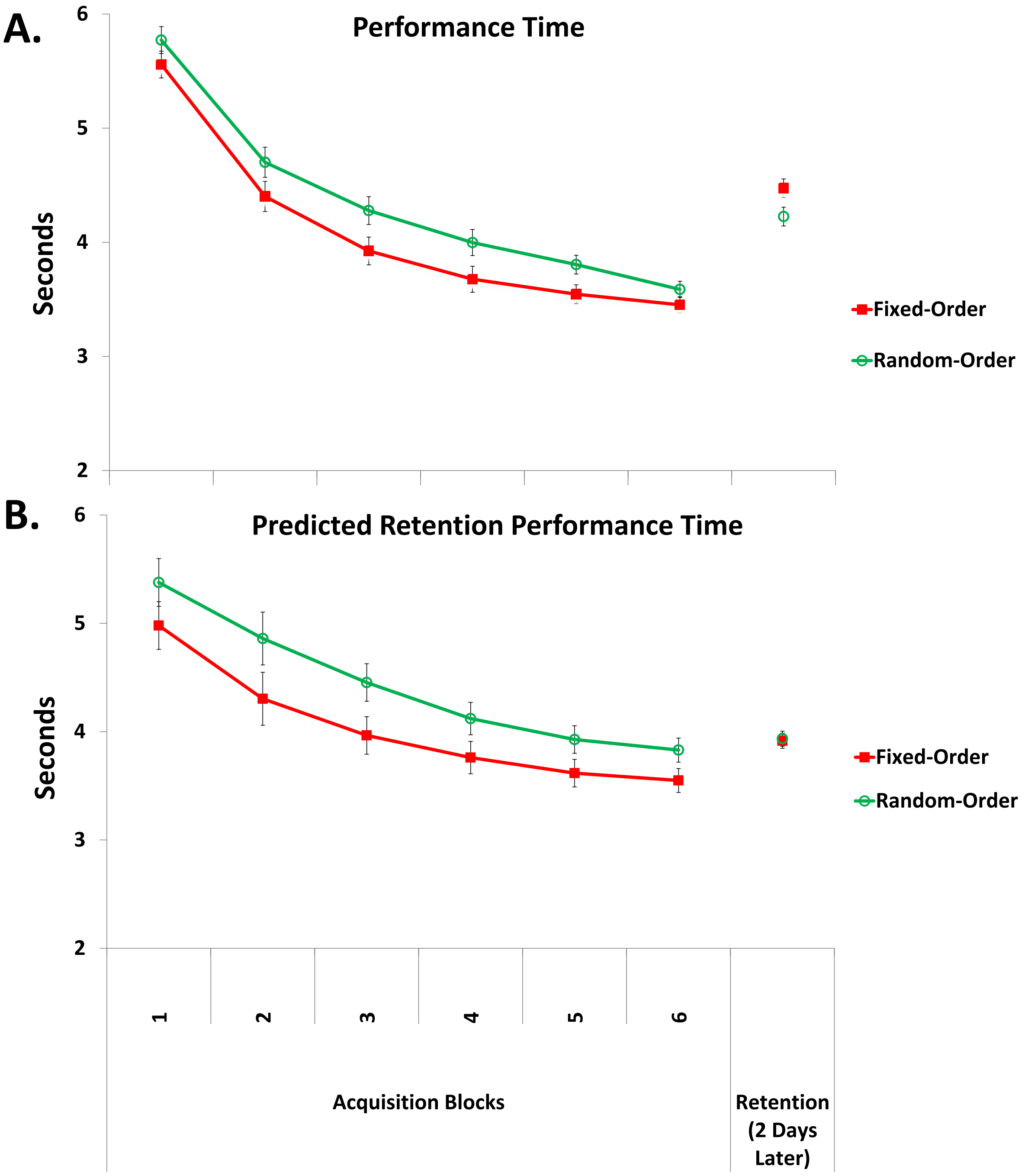

Figure 1. The mean performance time (Panel A) and predicted retention performance time (Panel B) as a function of fixed or random-order practice during acquisition. Error bars show the standard error of the difference between fixed and random-order practice.

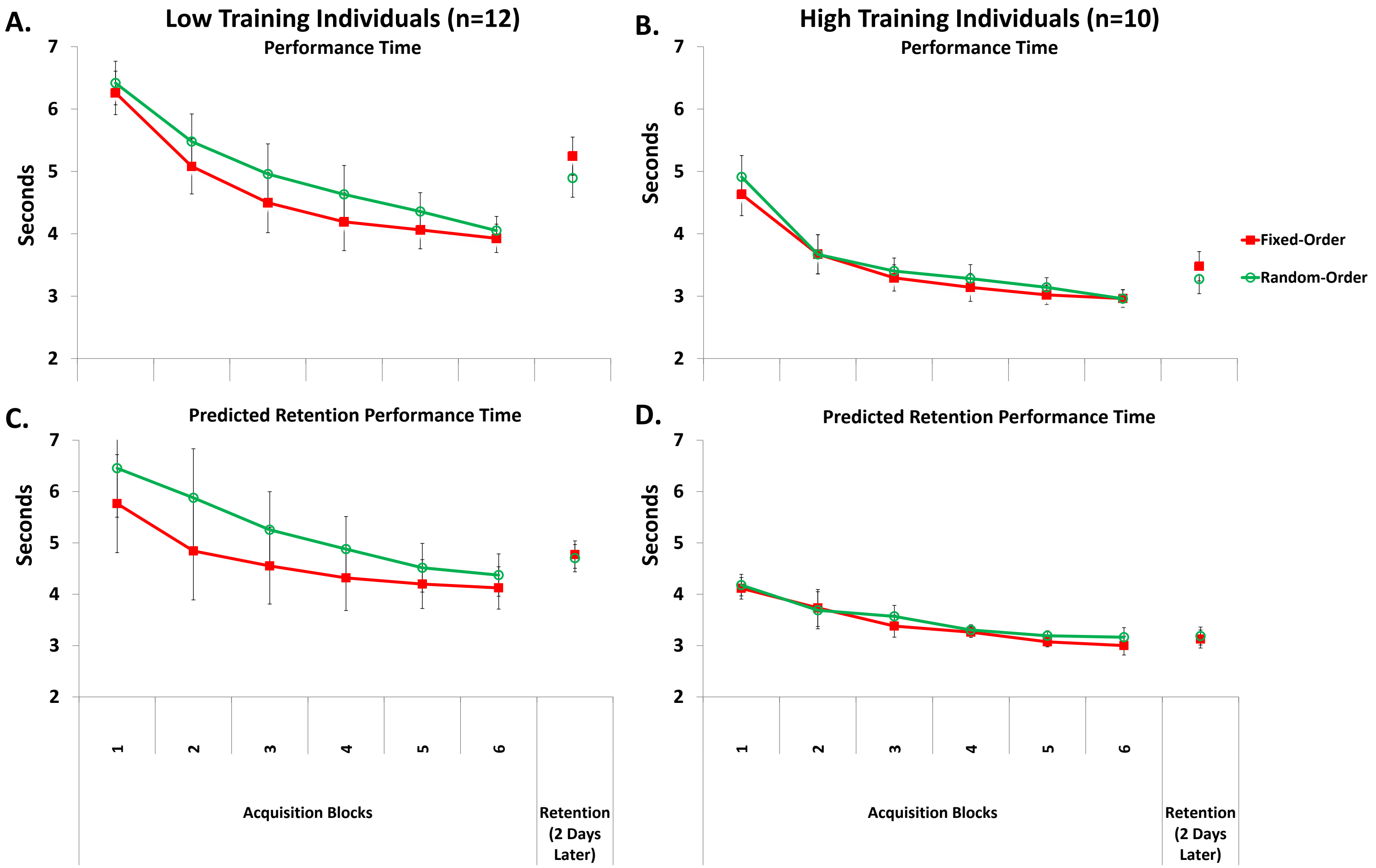

Figure 2. The mean performance time (panels A and B) and predicted retention performance time (panels C and D) as a function of training. Low Training Individuals had less than 9 years (the median) of formal training on piano, whereas High Training Individuals had more than 9 years. Error bars show the standard error of the difference between fixed and random-order practice.