# Efficiency of Blocked vs. Interleaved Multi-Session Training on Musical Style Recognition

Bhargav Kandlakuti

(Dr. Jessica Grahn, Rebekka Lagacé-Cusiac)

Western Centre for Brain and Mind, Western University, London, ON, Canada

Department of Psychology, Western University, London, ON, Canada

## Abstract

Category learning – the process of grouping items based on shared features – is essential for distinguishing elements like rhythm, composer, or style of a piece of music. While interleaved practice (alternating topics) is often linked to better retention, blocked practice (one topic at a time) is generally preferred. This study examines the effects of interleaved and blocked practice schedules on the ability to categorize musical styles of jazz pianists over multiple sessions. Participants completed interleaved and blocked practice, and then categorized novel clips based on the pianist they believed played the clip. No significant effect of practice schedule or test session delay was found. However, participants performed better when they had prior musical experience or when tested for memory on repeated clips. These findings suggest previous experience and type of recall may play a larger role than practice schedule on musical category learning, informing future curriculum design.


## Summary for Lay Audiences

Whether studying for a test or learning how to play new music, the way we schedule our practice matters. One common method is “blocked practice,” where one topic is focused on at a time. An alternative approach is “interleaved practice,” where different topics are mixed and learned simultaneously. An area that may benefit from practice scheduling adjustments is category learning – the process of grouping novel items or information based on shared characteristics. As such, category learning is essential in music where the ability to distinguish the rhythms or composer of a piece is essential. While research generally shows that interleaved practice can improve learning, it’s unclear how this applies to musical style category learning, especially over multiple days.

This study compared how blocked and interleaved practice affects participants' ability to recognize the musical styles of six jazz pianists. Participants first completed four training sessions each spaced one day apart, during which they listened to piano clips with the knowledge of who was playing them. Later, participants were tested on how well they could identify the pianist of new, unseen clips, during tests spaced one day and one week after practice. They also rated their confidence in their choices.

The study found no clear advantage for either blocked or interleaved practice when it came to recognizing musical styles. However, participants with at least five years of music experience performed better overall. People were also more accurate when categorizing clips they heard during training, rather than new clips. Interestingly, while participants overall felt more confident using the blocked schedule shortly after practice, this confidence did not reflect actual performance. This suggests that what feels easier to learn may not always lead to better results.

These findings are important for musical education and designing training programs. They suggest that the benefits of interleaved practice may not always apply, such as for tasks like distinguishing the musical styles of jazz pianists. Factors like prior experience may play a bigger role than practice schedule alone. Future research should continue exploring the most effective practice schedules for category learning to guide educational development.

**Keywords**

Interleaved practice; blocked practice; category learning; musical style; jazz pianists; musical education; random practice; contextual interference; auditory; retention.

**Acknowledgements**

I would like to express my sincere gratitude to my supervisor Dr. Jessica Grahn for the opportunity pursue this research endeavour and making this work possible. This project would not have been possible without her.

I am also extremely thankful to Rebekka Lagacé-Cusiac, for her invaluable guidance and support throughout this project. Thanks to her I gained several insights along the way and developed a stronger foundation of skills that were instrumental for the completion of this thesis.

My sincere thanks also go to Xinjie Dong, a partner on this project, for his collaboration and assistance throughout this research.

I extend my thanks to the Social Sciences and Humanities Research Council as well for funding this project and making it feasible.

I am also grateful to Dr. Zachary Easton for his teachings in the PhysPharm 4980 course. The lessons were exceptionally helpful both for the completion of this project and developing my knowledge base as a student.

## Table of Contents

## List of Tables

## List of Figures

# 1. Background

## 1.1 Interleaved and Blocked Practice Schedules

Optimizing learning strategies is a key aspect of teaching, occupational training, and skill acquisition in various disciplines including music. Modifying practice schedules or order is one way that this skill acquisition can be improved. Interleaved practice, a technique involving alternating between different topics or skills within a single training period, has been examined as a potential alternative to blocked learning, where individuals focus on learning one topic or skill before moving on to the next (Carvalho and Goldstone 2015). For instance, interleaved practice could be 123123123, whereas blocked learning would be 111222333, where each digit represents a different skill (Wong et al. 2020).

Several studies suggest that interleaved practice is more beneficial than blocked practice, particularly in the long term (Firth et al. 2021). This effect has been documented across a variety of fields and abilities, such as mathematics (Rohrer and Taylor 2007; Rohrer et al. 2015; Mielicki and Wiley 2022), painting (Kang and Pashler 2012), motor learning (Carson and Wiegand 1979), and sports techniques (Goode and Magill 1986). Studies have also been conducted comparing interleaved and blocked practice on musical ability such as quality and speed of a performance, noticing varying advantages (Abushanab and Bishara 2013; Carter and Grahn 2016; Stambaugh 2016).

## 1.2 Contextual Interference Effect

Interestingly, it has been proposed that facing challenges that reduce initial performance when learning can contribute to greater acquisition and retention of

information in the long term, a concept termed the “contextual interference effect” (Bjork 1994). Indeed, studies comparing interleaved and blocked practice schedules noticed that interleaved learning resulted in decreased performance during the learning stages itself, but greater performance later on (Firth et al. 2021; Mielicki and Wiley 2022). Accordingly, the studies also showed that blocked practice yielded greater benefits in the short term compared to the interleaved condition. For example, Rohrer and Taylor (2007) found that participants who were taught mathematics formulas performed better in the blocked condition immediately after their learning, while after a week had passed the interleaved group performed significantly more accurately.

The early benefits may also be the reason that blocked studying is usually found to be preferred and used more often by people over interleaved studying (Tauber et al. 2013; Carvalho and Goldstone 2015). A pivotal study by Kornell & Bjork (2008) noticed that when participants were taught to distinguish between the painting styles of various artists using both interleaved and blocked schedules, participants believed that blocked learning was superior despite performing better in the interleaved condition. Furthermore, this bias toward using blocked practice and believing it to be more effective than interleaved practice has been shown to extend to musical performance and musical style learning (Wong et al. 2020). For instance, in a study where clarinet players trained to perform musical pieces with both blocked and interleaved practice, they preferred the blocked condition despite performing better when interleaving (Carter and Grahn 2016). Previous findings suggest that this bias occurs due to a more fluid transition between learning material while blocking. When experiencing a piece of information or trying a new skill repetitively, it produces a higher sense of fluency as an

individual has multiple chances to learn it. However, interleaved learning precludes this repetition, instead alternating between pieces of information or types of skills, which induces a higher sense of difficulty in the individual and leads to a sense of inadequate learning (Kornell and Bjork 2008). Thus, the present study takes into account metacognitive judgements regarding the efficiency of blocked and interleaved practice, as participants may believe in the superiority of one condition over another despite the actual performance being the opposite.

## 1.3 Impact of Practice Schedules on Category Learning

One area of note that appears to be influenced by interleaved practice is category learning, the cognitive process through which we use prior experiences to classify objects or information (Deng and Sloutsky 2015). For example, category learning explains how a doctor may diagnose and classify a patient's illness based on their prior knowledge of symptoms related to that illness (Kornell and Bjork 2008). There have been studies comparing interleaved and blocked practice on category learning in many areas such as animals (e.g. recognizing the family that a bird belongs to; Tauber et al. 2013), distinguishing painter styles (Kornell and Bjork 2008), and the sciences (e.g. classifying chemical molecules; Eglington and Kang 2017).

Category learning is also crucial for developing musician skills such as aural skills, ear training, and overall music perception ability. The aforementioned skills are regarded as crucial across musical education, and the ability to discriminate between composers and musicians, detect cultural and era-specific influences, and recognize the characteristics of a musical excerpt are some key considerations in musical curricula. However, there are limited studies assessing the impact of interleaved practice on

category learning in the field of music, such as for developing the ability to discriminate between the composer, era, or instruments of a piece of music (Carter and Grahn 2016; Wong et al. 2020).

Wong et al. (2020) conducted a study comparing interleaved and blocked practice on a form of musical category learning, specifically the ability of participants to recognize and differentiate the musical styles of different composers. To note, a musical style can be considered as the characteristics that constitute the composition and sound of a musical piece (Baroni 2009). Non-musician participants (those with less than 4 years of musical experience) were given music clips of 12 composers and asked to learn half in an interleaved fashion and the other half in a blocked fashion. Following a 15-second task to distract participants, they underwent a testing phase involving listening to unfamiliar clips written by the same composers and matching them to the musician who they believed composed the clip. Results from the study showed that participants more effectively matched a novel clip to the correct composer in the interleaved condition. However, while this study supports the use of interleaved practice for category learning in music, the testing and practice phases were both conducted on the same day, and there is a lack of research comparing these practice schedules over several days for musical learning (Carter and Grahn 2016; Wong et al. 2020). Testing learning acquisition in the long-term is also particularly significant as some studies have shown that the benefits of interleaved practice are more noticeable after a period of delay (Goode and Magill 1986; Rohrer and Taylor 2007; Taylor and Rohrer 2010; Stambaugh 2016; Mielicki and Wiley 2022)

### 1.4 Present Study

This study aims to evaluate the effectiveness of interleaved versus blocked practice in the context of category learning of musical styles, using a multi-session design including four practice sessions and two delayed testing sessions. Because most formal education occurs over multiple days with repeated exposure, this extended schedule better reflects real-world learning environments. Building on previous research that explored how practice schedules affect the ability to recognize classical composers' styles, this study aimed to examine the effects of practice schedules on recognizing the musical styles of jazz improviser pianists. This study also includes both musicians and non-musicians, addressing a limitation in Wong et al. (2020), which only tested non-musicians. Prior research suggests that interleaved practice may benefit novices even when they are learning complex skills, making it relevant to test both those with and without prior experience (Ollis et al. 2005). Based on previous evidence supporting interleaved practice over blocked practice for category learning, it is hypothesized that interleaved practice will be more effective than interleaved practice. More specifically, participants will more accurately distinguish jazz improviser pianist musical styles both one day and one week after completing their practice.

## 2. Materials and Methods

### 2.1 Participants

32 students from Western University were recruited to participate in this study through mass email and verbal recruitment. Upon recruitment, participants were classified as musicians (over 5 years of musical education) or non-musicians (less than

5 years of musical education). A target sample of 24 participants was determined based on power analysis. For a within-participant study with an effect size of 0.4 (dz=0.4, medium effect size) and a power of 0.8, the analysis showed that a sample of 24 participants is required to detect a significant effect between the treatment groups. To account for participant dropout, we aimed to recruit additional participants obtaining 32 in total. Participants who dropped out of the study or had not completed all sessions were excluded from data analysis, leading to a final sample of 17 participants. The final sample consisted of 12 musicians and 5 non-musicians, 9 females and 8 males, and participants were between the ages of 18 to 21 ($M$ = 19.58)

Participants were compensated $10 per hour for their involvement in the study, up to $30 for completing all six sessions. The Western University Non-Medical REB at Western University gave ethics approval for this study and informed written consent was obtained from participants.

## 2.2 Musical Stimuli

To create the experimental task, 6 jazz improviser pianists were selected to obtain clips from: Brad Mehldau, Bill Evans, Chick Corea, Thelonious Monk, Oscar Peterson, and Marian McPartland. 56 clips were drawn from each pianist, 10 seconds long each. To reduce the chances that participants were already familiar with the musical styles of the pianists, songs that were not relatively well-known were selected, listed in Supplementary Table 1. Participants were to listen to these clips to try to learn to recognize and distinguish between the musical playing styles of each of the 6 jazz improviser pianists. Clips that contained only piano solos and no other background music were selected to prevent overlap with other instruments or sounds. Furthermore,

clips featuring musical improvisation were selected to ensure that the differences between individual pianists were sufficiently pronounced and perceptible. Unlike performances of notated music, improvisation involves performers expressing their unique musical playing style -- such as variations in rhythm, harmony, phrasing, and emotional tone. The clips were edited and normalized using Audacity and then cross-checked with two expert musicians to ensure that they were a piano solo and contained improvisation. At most, two 10-second clips were taken from a given song around the first and second-thirds of a song, to minimize the quantity of repeated elements between clips. If a clip did not contain sufficient improvisation or had too little piano playing time, it was replaced with a clip taken at a similar time interval. Once completed, this experimental task was facilitated online through Pavlovia.

## 2.3 Experimental Task Design

The experiment consisted of six total sessions outlined in Fig. 1. The first four sessions were practice sessions, while the fifth and sixth sessions were testing sessions. Sessions 1, 5, and 6 were conducted in person, whereas Sessions 2 through 4 were completed online. The goal of the practice sessions was to teach participants to recognize and categorize the musical styles of each jazz pianist, which they would later be tested on during the fifth and sixth sessions. Each practice section consisted of two halves, one presented in a blocked format and the other in an interleaved format, for which the jazz pianists were split into two groups 1) Bill Evans, Chick Corea, Oscar Peterson, and 2) Brad Mehldau, Thelonious Monk, Marian McPartland. To ensure that there was no effect based on these grouping decisions, the sequence of learning was counterbalanced across participants. If one participant experienced Group 1 in the

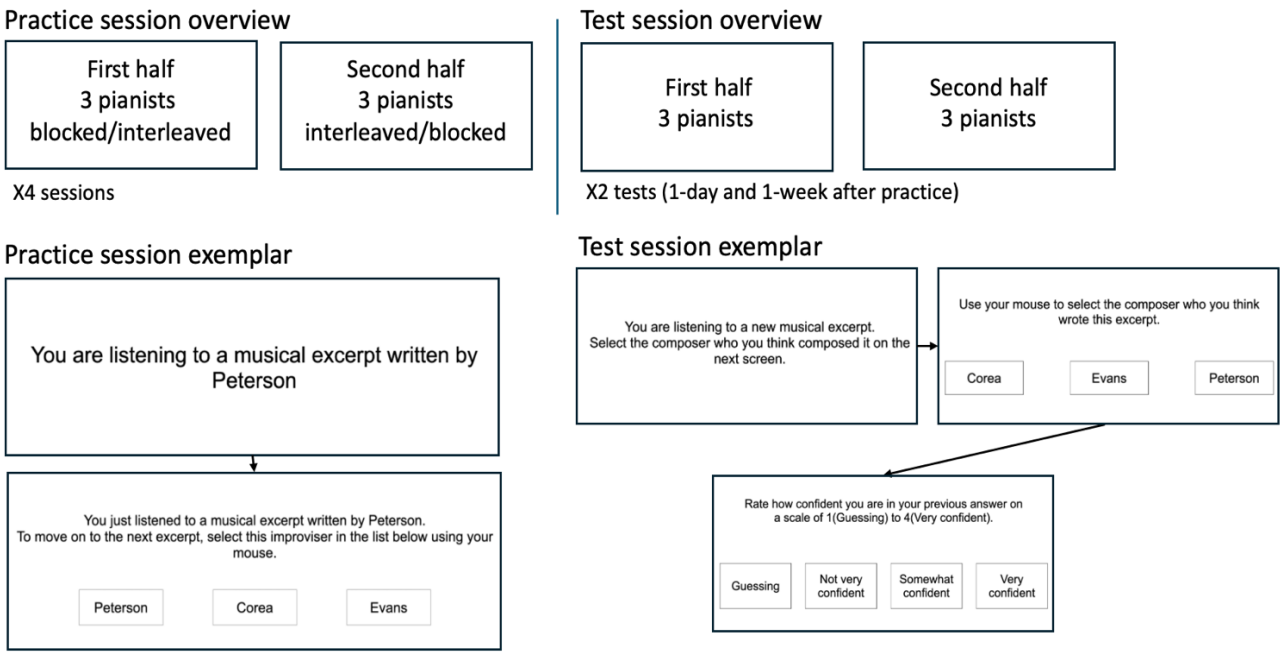



**Figure 1. Experimental Task Overview**

Practice and test session overviews highlight the broad sequence of events occurring over the six sessions (four practice sessions followed by two test sessions). Practice sessions are composed of two halves, where participants learn the musical styles of three jazz improviser pianists in each half. When one half was completed in a blocked or interleaved order, the next half would be completed in the opposite form. During test sessions, participants were tested on their ability to recognize the musical styles of jazz improviser pianists from practice sessions. Practice and test session exemplars display example trials that a participant would see during each session. During practice sessions, participants listened to ten second clips to learn the musical styles of the jazz pianists. During test sessions, participants listened to a clip and were asked to select who they believed played the clip.

interleaved format and Group 2 in the blocked format, the next participant received the opposite configuration. Additionally, the order in which participants learned each group was also counterbalanced. This resulted in four different counterbalance conditions, with every set of four participants assigned to a different sequence.

During the first practice session, participants listened to 14 clips from each pianist, for a total of 84 clips. When participants were learning in a blocked order, they listened to 14 clips from each pianist before moving on to the next for each of the three jazz pianists. When participants were completing the interleaved practice, they listened to 14 clips from each pianist in a mixed order without ever repeating a pianist's clips more than once at a time. During both these halves, participants were told the name of the pianist on the screen while the clip was playing. To ensure participants were actively engaging with practice, they were then asked to select the pianist who played the clip from a list of three pianists. Following the completion of the interleaved and blocked halves, the first session was done. The next three practice sessions were aimed to be completed roughly 1 day (up to 24-48 hours) apart and were delivered online consisting of the same task. However, rather than listening to 14 novel clips from each pianist, 4 clips were repeated clips from the first session, while only 10 clips were novel.

After completion of the four practice sessions, participants were asked to attend in person for the fifth and sixth sessions – the testing sessions. The fifth session was placed 1 day after the conclusion of the practice sessions while the sixth session took place 1 week after the fifth session to test the short and long-term effects of the practice schedule on learning. During the testing sessions, participants listened to 10 clips from each composer (total 60), of which 4 were repeated clips from the practice sessions

(testing memory of clips), while 6 were novel clips that the participants had not yet heard (testing transfer of knowledge – the ability to apply learning to previously unseen stimuli). Participants listened to these clips in a random order without knowing who played them, and were then asked to select who they believed played the clip from a list of 3 pianists (3 options based on the pianist groups from the practice sessions). Accuracy was measured here as the percentage of correctly categorized clips. Subsequently, after each response, participants were asked to rate their confidence in their response on a scale of 1 (Guessing) – 4 (Very confident), with options including “Guessing”, “Not very confident”, “Somewhat confident”, to “Very confident”. Measures of confidence were used to assess participants’ metacognitive judgements regarding preferences regarding practice schedule. After the sixth and final session, participants completed a questionnaire assessing demographics.

## 2.4 Data Analysis

### 2.4.1 Participant Accuracy

Data was first checked to confirm a normal distribution, and that variance and covariance were consistent across groups. A four-way 2x2x2x2 mixed ANOVA was used to analyze the influences of interleaved and blocked practice schedules on musical style learning. The independent variables were practice schedule (interleaved or blocked), test session delay (1 day or 1 week; corresponding to sessions 5 and 6 respectively), repetition (repeated clips or novel clips; indicates whether clips were repeated from practice or entirely new), and musicianship status (musician or non-musician). Accuracy (%) was measured as the dependent variable. Within-subject factors were practice schedule, test session delay, and repetition, whereas the between-subject factor was

musicianship status. Tukey's HSD post hoc analysis was conducted when main effects or interactions were found. Data was analyzed in RStudio version 2024.12.1+563 using the emmeans and afex packages.

#### 2.4.2 Confidence Ratings

Confidence ratings analysis was done with the same method used to analyze participant accuracy. A four-way 2x2x2x2 mixed ANOVA with the same independent variables was conducted, while confidence rating (scaled 1-4) was measured as the dependent variable. Within-subject and between-subject factors were the same as when measuring accuracy. Tukey's HSD post hoc analysis was conducted when main effects or interactions were found. Data was analyzed in RStudio version 2024.12.1+563 with the emmeans and afex packages.

# 3. Results

## 3.1 Impact of practice schedule and test session delay on participant accuracy when categorizing jazz improviser pianists

To assess accuracy, the proportion of pianists that participants correctly categorized was measured and expressed as a percent (Fig. 2). No main effect was found for practice schedule ($F(1,15) = 0.07$, $p = 0.79$) and test session delay ($F(1,15) = 0.18$, $p = 0.68$). There was a significant main effect for musicianship status, $F(1,15) = 7.36$, $p<0.05$, $\eta^2g = 0.14$ and repetition $F(1,15) = 6.04$, $p<0.05$, $\eta^2g = 0.03$, where musicians had a higher accuracy than non-musicians and participants performed better on repeated clips. No interactions were found.

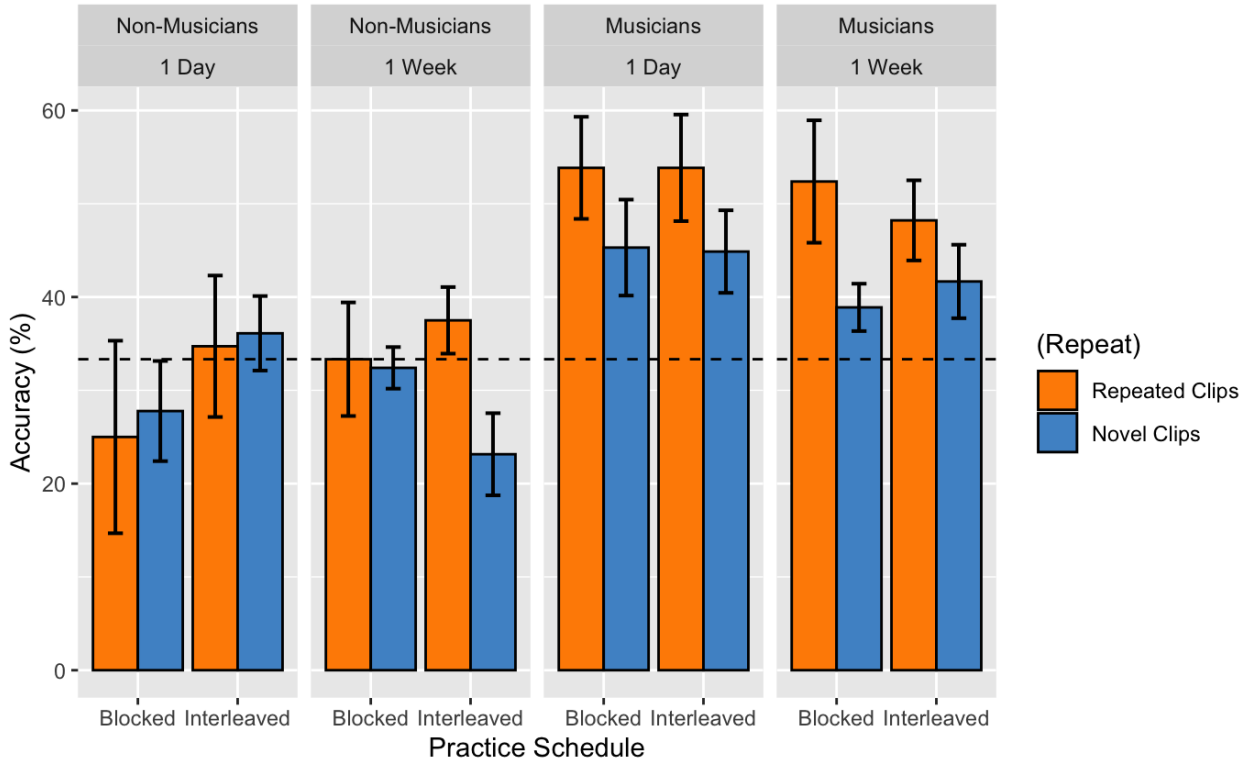
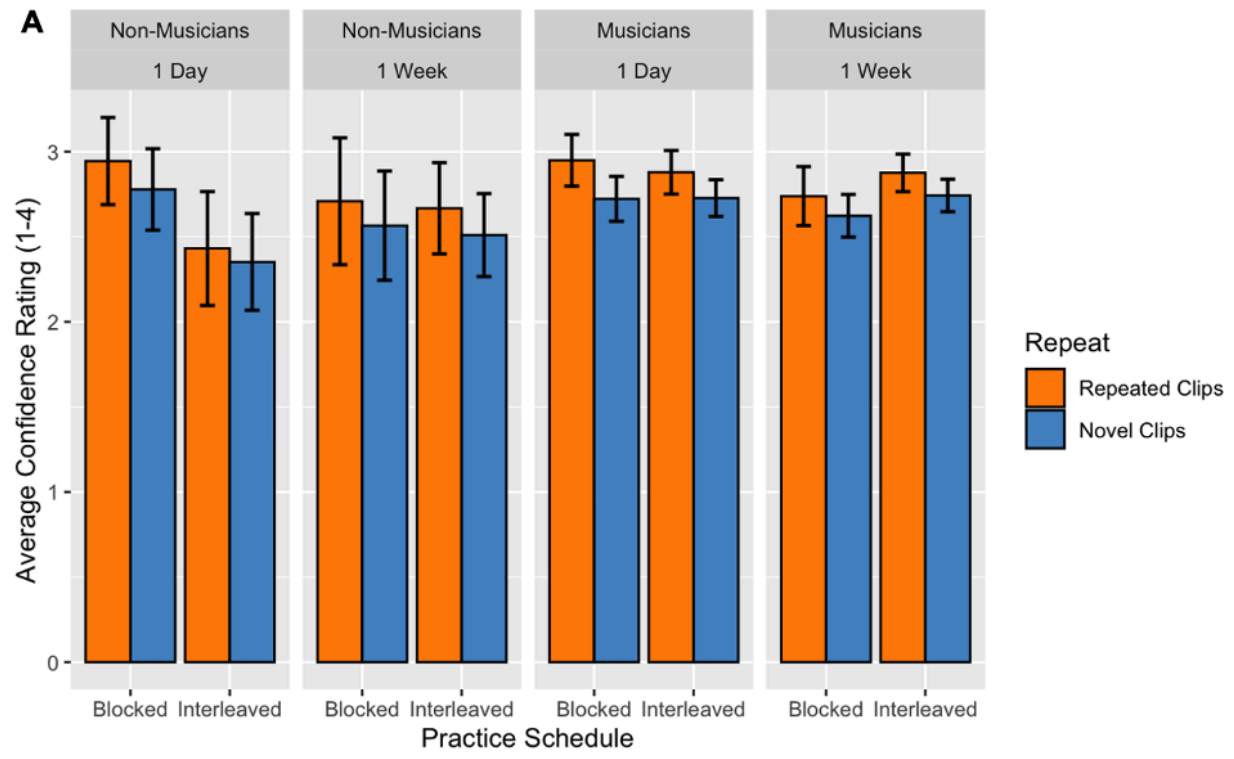



**Figure 2. Mean Accuracy on Categorizing Jazz Improviser Pianists in Blocked and Interleaved Practice Schedules Across Treatment Groups**

Mean interleaved and blocked practice schedule effects on categorizing jazz improviser pianists, measured as accuracy (%). Participants first completed four practice sessions spaced one day apart to learn jazz improviser pianist music styles. The tested jazz pianists included Brad Mehldau, Bill Evans, Chick Corea, Thelonious Monk, Oscar Peterson, and Marian McPartland. Following practice, participants then completed testing sessions spaced 1 day and 1 week apart (measuring effects of test session delay). During testing sessions, participants were given 10 second clips and asked to select the pianist that played the clip. Musician (n= 12) and non-musician (n= 5) participants categorized clips that were either repeated from practice or novel clips to their respective pianist, and accuracy on their responses was measured. There was no main effect of practice schedule (blocked vs. interleaved) or test session delay (1 day vs. 1 week). There was a main effect of whether clips were repeated from practice and musicianship status, where old clips and musician status resulted in higher accuracy ($p<0.05$). Four-way 2x2x2x2 Mixed Anova with Tukey's HSD post hoc test was conducted to assess significant differences. Dashed line represents expected accuracy from random chance responses alone. Error bars represent SE.

## 3.2 Impact of practice schedule and test session delay on participant confidence ratings when categorizing jazz improviser pianists

To assess confidence, participant's mean confidence ratings across responses was measured on a scale of 1-4 (Fig. 3). There was no main effect of practice schedule ($F(1,15) = 2.19$, $p = 0.16$) or test session delay ($F(1,15) = 0.48$, $p = 0.50$). A significant main effect was found for repetition $F(1,15) = 10.22$, $p<0.01$, $\eta^2g = 0.01$, where participants were more confident on repeated clips. There was also a significant interaction between practice schedule and test session delay, $F(1,15) = 8.11$, $p<0.05$, $\eta^2g = 0.01$, where participants were more confident in the blocked order after the 1-day test. No other main effects or interactions were found.


A
Non-Musicians
1 Day
Non-Musicians
1 Week
Musicians
1 Day
Musicians
1 Week
Average Confidence Rating (1-4)
0
1
2
3
Blocked
Interleaved
Practice Schedule
Repeat
Repeated Clips
Novel Clips

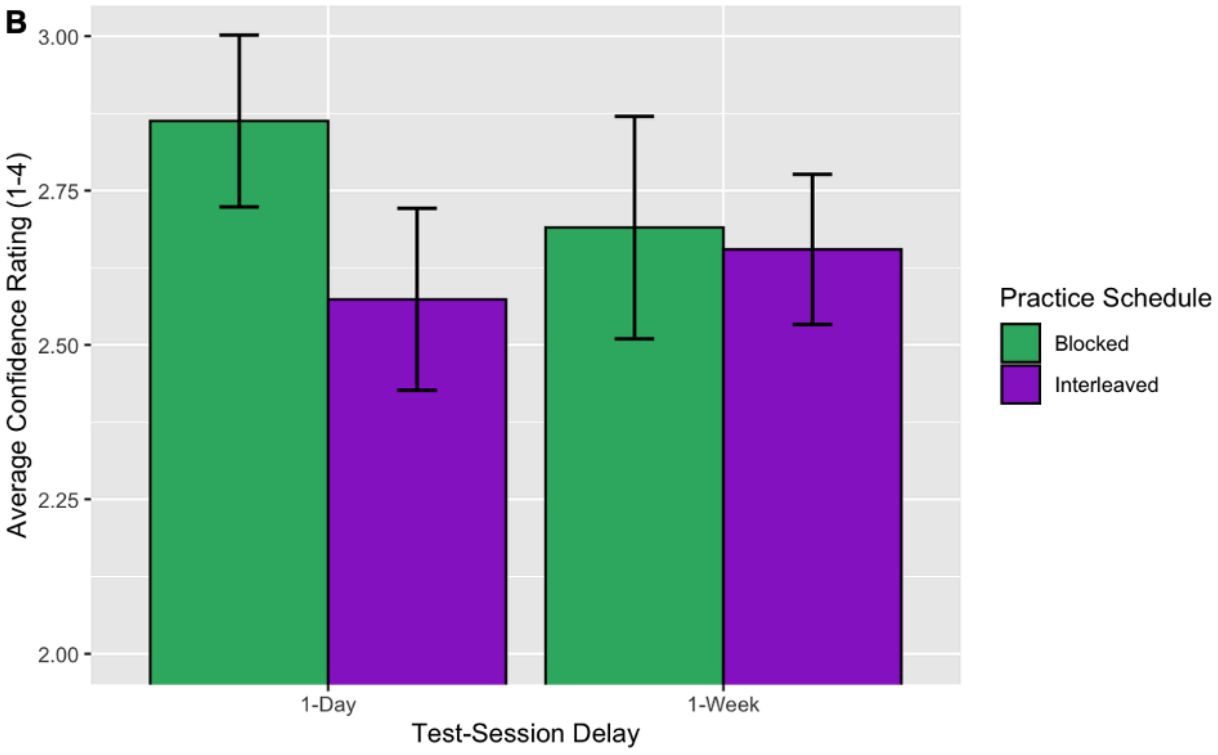

B
Average Confidence Rating (1-4)
2.00
2.25
2.50
2.75
3.00
1-Day
1-Week
Test-Session Delay
Practice Schedule
Blocked
Interleaved

**Figure 3. Mean Confidence Ratings in Interleaved and Blocked Practice Schedules Across Treatment Groups**

**A)** Practice schedule effects on metacognitive judgements of accuracy, measured as the mean confidence rating. Participants first completed four practice sessions spaced one day apart to learn jazz improviser pianist music styles. The tested jazz pianists included Brad Mehldau, Bill Evans, Chick Corea, Thelonious Monk, Oscar Peterson, and Marian McPartland. Following practice, participants then completed testing sessions spaced 1 day and 1 week apart (measuring effects of test session delay). During testing sessions, participants were given 10 second clips and asked to select the pianist that played the clip. Clips were either repeated from practice or novel clips. Musician (n= 12) and non-musician (n= 5) participants were then asked to rate their confidence in their responses following each trial. Ratings were on a scale of 1-4, with options ranging from “Guessing” (1), “Not very confident” (2), “Somewhat confident” (3), to “Very confident” (4). There was no main effect of practice schedule (blocked vs. interleaved), test session delay (1 day vs. 1 week), or musicianship status (musician vs. non-musician). There was a main effect of whether clips were repeated from practice, where participants rated higher confidence on old clips ($p<0.01$). **B)** Interaction was found between practice schedule and test session delay, where participants were more confident during the test session taken after 1 day when using a blocked practice schedule ($p<0.05$). Four-way 2x2x2x2 Mixed Anova with Tukey’s HSD post hoc test was conducted to assess significant differences. Dashed line represents expected accuracy from chance responses alone. Error bars represent SE.

# 4. Discussion

The present study sought to investigate the impacts of blocked and interleaved practice schedules on musical category learning over multiple sessions. To do so, participants' ability to recognize and distinguish between the musical styles of six jazz improviser pianists was used as a novel measure of category learning in music. Based on prior studies testing category learning, it was hypothesized that interleaved practice would result in better learning of the jazz pianists' musical styles, both in the short and long term. However, no significant effect of practice schedule was found in the tests conducted 1 day or 1 week after completion of the practice phase, disagreeing with the hypothesis that interleaved practice would be superior. These findings contradict past studies that show benefits of interleaved practice over blocked practice in areas such as mathematics (Rohrer and Taylor 2007; Rohrer et al. 2015; Mielicki and Wiley 2022) motor skills (Carson and Wiegand 1979), learning musical pieces (Carter and Grahn 2016), learning to identify musical intervals (Wong et al. 2021), and musical category learning (Wong et al. 2020). Furthermore, as opposed to the results found in this study, studies have shown benefits of interleaved practice in the long term (Taylor and Rohrer 2010; Stambaugh 2016; Firth et al. 2021; Mielicki and Wiley 2022).

Although the findings of the current study diverge from many existing literature, the data remain valid. Participants clearly demonstrated learning over the four practice sessions, evidenced by significantly higher accuracy when categorizing jazz pianists using clips that had been repeated from the practice sessions shown in Fig. 1. This trend reflects greater memorization than transfer of knowledge to new clips. Transfer of knowledge, which can be defined as the ability to categorize or classify novel stimuli

based on past knowledge, is generally more difficult than memorization that involves categorizing older information (Salomon and Perkins 1989; Firth et al. 2021). As such, participants performing better when tested for memorization than transfer of knowledge indicates that they were successfully remembering clips from the practice sessions, but this learning did not lead to any significant difference between interleaved and blocked practice. Therefore, while the study design is still valid, it does not entirely support the null hypothesis of no difference between interleaved and blocked practice on musical category learning, as the findings are specific to jazz pianists and may not be generalizable across broader musical category learning. One other limitation is that the sample size after exclusion was only 17, so future studies with more participants are needed to build on these results.

While studies generally show an advantage of interleaved practice (Firth et al. 2021), these findings are not always consistent. For example, when comparing interleaved and blocked practice for learning to recognize musical intervals (differences in pitch between two notes), Wong et al. (2021) found that interleaving only conferred an advantage under certain circumstances. When participants had reference information, interleaved learning was more beneficial, but when the reference information was not present, there was no difference in practice schedules. Likewise, in another study that manipulated practice scheduling while teaching participants to distinguish non-musical sounds, there was no difference between interleaving and blocking (Obasih et al. 2023). In contrast, Wong et al. (2020) found that interleaved practice improved participant's ability to distinguish composers' musical styles across four eras. Unlike the present study, Wong et al. (2020) conducted a single practice and

training session on the same day, spaced by a 15-second distractor task. Thus, the benefits of interleaved practice may not extend to learning that takes place over multiple days when it comes to musical category learning. Comparing past literature to the present findings suggests that the interleaving benefit may not be universal. That is, the advantages of interleaved schedules may be specific to certain conditions or forms of category learning, such as comparing composers across eras.

Another reason the current findings may differ from previous research is the potential difficulty of the task itself. Recognizing and accurately classifying the musical styles of jazz improvisers may have been too challenging. As shown in Fig. 1, non-musicians in particular often performed below random chance level. In this study, participants studied 10-second clips from pianists. Learning from longer clips or potentially entire songs from each pianist could reduce task difficulty lead to different results. Similarly, increasing the number of practice sessions, to eight or ten for example, could provide a more sufficient learning experience. Indeed, studies have shown spreading sessions over longer periods and adding more time between practice sessions can improve learning (Smolen et al. 2016). Another limitation of this study was the use of fixed jazz pianist groups. Randomly assigning the six pianists into groups of three or choosing entirely different jazz pianists could be a step for future research.

## 4.1 Influences of Prior Experience on Musical Category Learning

The current study found that participants learned the musical styles of jazz improviser pianists more effectively when they had more than 5 years of music experience, but there was no interaction of musicianship with practice schedule (Fig. 1). This relationship between previous experience and practice schedule is not entirely

clear in other studies either. For example, where Brunmair and Richter (2019) found that previous experience resulted in a greater benefit of interleaved practice, Carpenter and Mueller (2013) noticed that blocked practice was actually more useful in helping novices learn to pronounce words from unfamiliar languages. Thus, there is no clear effect of expertise on practice schedule efficacy.

The observation that musicians performed better than non-musicians at categorizing jazz improviser pianists is likely due to their techniques acquired through musical education. However, musicians may have also been more interested in the experimental task to begin with, leading to greater focus, learning, and retention of information. Obtaining data on participants' engagement during practice sessions could help clarify how factors such as motivation influence musical category learning.

## 4.2 Metacognitive Judgements and Confidence Ratings

Participants' metacognitive judgements were measured through confidence ratings collected after each jazz pianist identification task (Fig. 2). Although prior research suggests that individuals often prefer blocked practice and believe it to be more effective than interleaved practice, the current study found no overall effect of practice schedule on confidence ratings. Instead, a significant interaction was found between practice schedule and test session delay, where participants were more confident in the blocked condition over the interleaved condition one day after practice, but this difference was absent in the test session conducted one week later. These findings partially align with previous studies showing that individuals preferred blocked practice in tasks such as distinguishing between painting styles (Kornell and Bjork 2008), learning clarinet pieces (Carter and Grahn 2016), or differentiating between

composer musical styles (Wong et al. 2020). However, unlike these studies, the present study did not show any performance advantage for either practice schedule, despite participants believing in the superiority of the blocked order one day after practice. Looking further into the current data, participants were more confident in categorizing clips that were repeated from practice rather than novel clips, which aligned with their actual performance accuracy as well (Fig. 1). This indicates that participants were at least partially able to gauge their accuracy, supporting the relevance of the confidence ratings as a measure of metacognitive judgement in this study.

The disconnect between participants' preferred practice schedule and the actual effectiveness of practice schedules occurs often (Tauber et al. 2013; Carvalho and Goldstone 2015). A potential reason that participants preferred blocked practice in this study, at least during the test session conducted one day after practice, may be due to a sense of familiarity. Most individuals utilize blocked practice when attempting to learn new skills, and it is typically the preferred schedule used in educational practices. Thus, participants may have had a bias towards a practice schedule that they are familiar with using. Furthermore, as indicated in past research, interleaved learning can be challenging due to the difficulty of attempting to process several categories or skills at once. This difficulty may have caused participants to feel that they were not learning as effectively under the interleaved condition, leading to a preference for blocked practice (Kornell and Bjork 2008; Carter and Grahn 2016).

# 5. Conclusion

In conclusion, this study did not support the hypothesis that interleaved practice was more beneficial than blocked practice for jazz pianist musical style category learning. Furthermore, these findings persisted over testing sessions completed one day and one week after practice. Instead, participants showed greater category learning when they had prior music experience and when tested for memory rather than transfer of knowledge. There is a lack of research on the impact of practice schedules on category learning over multiple days, making these findings useful for guiding future educational practices. As most formal education involves repeated exposure over time, future studies should continue to explore the impact of different timelines -- such as several weeks or months -- on interleaved practice. Additionally, the findings of the present study are limited to six jazz improviser pianists. To broaden these data, future research should assess categorization of different pianists or forms of musical category learning, such as comparing instruments across various cultures. This study suggests that interleaved practice may not always be more beneficial than blocked practice and instead may depend on the forms of category learning and timelines used. For instance, there may be no impact of practice schedule on learning to distinguish jazz pianist musical styles. Comparing participant confidence ratings showed that participants partially preferred the blocked condition over the interleaved, which aligns with past research. However, as there was no actual benefit of either schedule, it suggests that individual learning preferences do not always reflect actual learning effectiveness. Continued research comparing practice schedules on musical style category learning should be conducted to continue to develop optimized educational practices.

## 6. References

Abushanab B, Bishara AJ. 2013. Memory and metacognition for piano melodies: Illusory advantages of fixed- over random-order practice. Mem Cognit. 41(6):928–937. doi:10.3758/s13421-013-0311-z.

Baroni M. 2009. A different kind of similarity: The recognition of style in listening. Musicae Scientiae. 13(1_suppl):119–138. doi:10.1177/102986490901300106.

Brunmair M, Richter T. 2019. Similarity matters: A meta-analysis of interleaved learning and its moderators. Psychol Bull. 145(11):1029–1052. doi:10.1037/bul0000209.

Carpenter SK, Mueller FE. 2013. The effects of interleaving versus blocking on foreign language pronunciation learning. Mem Cognit. 41(5):671–682. doi:10.3758/s13421-012-0291-4.

Carson LM, Wiegand RL. 1979. Motor Schema Formation and Retention in Young Children. J Mot Behav. 11(4):247–251. doi:10.1080/00222895.1979.10735193.

Carter CE, Grahn JA. 2016. Optimizing Music Learning: Exploring How Blocked and Interleaved Practice Schedules Affect Advanced Performance. Front Psychol. 7. doi:10.3389/fpsyg.2016.01251.

Carvalho PF, Goldstone RL. 2015. The benefits of interleaved and blocked study: Different tasks benefit from different schedules of study. Psychon Bull Rev. 22:281–288. doi:10.3758/s13423-014-0676-4.

Deng W (Sophia), Sloutsky VM. 2015. The development of categorization: Effects of classification and inference training on category representation. Dev Psychol. 51(3):392–405. doi:10.1037/a0038749.

Eglington LG, Kang SHK. 2017. Interleaved presentation benefits science category learning. J Appl Res Mem Cogn. 6(4):475–485. doi:10.1016/j.jarmac.2017.07.005.

Firth J, Rivers I, Boyle J. 2021. A systematic review of interleaving as a concept learning strategy. Review of Education. 9(2):642–684. doi:10.1002/rev3.3266.

Goode S, Magill RA. 1986. Contextual Interference Effects in Learning Three Badminton Serves. Res Q Exerc Sport. 57(4):308–314. doi:10.1080/02701367.1986.10608091.

Kang SHK, Pashler H. 2012. Learning Painting Styles: Spacing is Advantageous when it Promotes Discriminative Contrast. Appl Cogn Psychol. 26(1):97–103. doi:10.1002/acp.1801.

Kornell N, Bjork RA. 2008. Learning Concepts and Categories. Psychol Sci. 19(6):585–592. doi:10.1111/j.1467-9280.2008.02127.x.

Mielicki MK, Wiley J. 2022. Exploring the necessary conditions for observing interleaved practice benefits in math learning. Learn Instr. 80:101583. doi:10.1016/j.learninstruc.2022.101583.

Obasih CO, Luthra S, Dick F, Holt LL. 2023. Auditory category learning is robust across training regimes. Cognition. 237:105467. doi:10.1016/j.cognition.2023.105467.

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(3):229–244. doi:10.1016/j.actpsy.2004.08.003.

Rohrer D, Dedrick RF, Stershic S. 2015. Interleaved practice improves mathematics learning. J Educ Psychol. 107(3):900–908. doi:10.1037/edu0000001.

Rohrer D, Taylor K. 2007. The shuffling of mathematics problems improves learning. Instr Sci. 35(6):481–498. doi:10.1007/s11251-007-9015-8.

Salomon G, Perkins DN. 1989. Rocky Roads to Transfer: Rethinking Mechanism of a Neglected Phenomenon. Educ Psychol. 24(2):113–142. doi:10.1207/s15326985ep2402_1.

Smolen P, Zhang Y, Byrne JH. 2016. The right time to learn: mechanisms and optimization of spaced learning. Nat Rev Neurosci. 17(2):77–88. doi:10.1038/nrn.2015.18.

Stambaugh LA. 2016. Implications of extrinsic cognitive load on three levels of adult woodwind players. Psychol Music. 44(6):1318–1330. doi:10.1177/0305735615627206.

Tauber SK, Dunlosky J, Rawson KA, Wahlheim CN, Jacoby LL. 2013. Self-regulated learning of a natural category: Do people interleave or block exemplars during study? Psychon Bull Rev. 20(2):356–363. doi:10.3758/s13423-012-0319-6. http://link.springer.com/10.3758/s13423-012-0319-6.

Taylor K, Rohrer D. 2010. The effects of interleaved practice. Appl Cogn Psychol. 24(6):837–848. doi:10.1002/acp.1598.

Wong SSH, Chen S, Lim SWH. 2021. Learning melodic musical intervals: To block or to interleave? Psychol Music. 49(4):1027–1046. doi:10.1177/0305735620922595.

Wong SSH, Low ACM, Kang SHK, Lim SWH. 2020. Learning Music Composers' Styles: To Block or to Interleave? Journal of Research in Music Education. 68(2):156–174. doi:10.1177/0022429420908312.

# 7. Supplementary Information

**Supplementary Table 1:** Song List Used to Generate Experimental Clips

| Pianist | Song |
|---|---|
| Corea | Monk's Dream (Live) |
| Corea | Blue Monk (Live) |
| Corea | Waltz for Debby |
| Corea | Pastime Paradise |
| Corea | Brasilia |
| Corea | Thinking Of You (Live) |
| Corea | Oblivion |
| Corea | Armando's Rhumba (Live) |
| Corea | Dusk In Sandi (Live) |
| Corea | It could happen to you (live) |
| Corea | Portrait #1 - Vilnius |
| Corea | April Snow (Live) |
| Corea | Brasilia |
| Corea | Jonathan Livingston Seagull |
| Corea | This Nearly Was Mine |
| Corea | Livestream Highlights - "You’re Everything" |
| Corea | Prelude Op. 11, No. 2 (Solo Piano 1983) |
| Corea | The Falcon (Live) |
| Corea | Swedish Landscape |
| Corea | But Beautiful (Live) |
| Corea | Ask Me Now (Live) |
| Corea | Yesterdays (Official Audio) |
| Corea | Chick Practicing "The One Step" |
| Corea | So In Love (Live) |
| Corea | How Deep Is The Ocean? (Live) |
| Corea | The Yellow Nimbus (Live) |
| Corea | Song For Sally (1971) |
| Corea | Piano Improvisations Vol.1 - Where Are You Now? - Picture 1 |
| Evans | All The Things you are |
| Evans | A Time For Love |
| Evans | Midnight Mood |
| Evans | On A Clear Day (You Can See Forever) |
| Evans | Never Let Me Go |
| Evans | Love Is Here To Stay |
| Evans | When I Fall In Love |
| Evans | Everything Happens To Me |

| | |
|---|---|
| Evans | April In Paris |
| Evans | I Loves You, Porgy |
| Evans | What Kind Of Fool Am I? |
| Evans | Ornithology |
| Evans | Here's That Rainy Day (Verve Records 1968) |
| Evans | Medley: My Favorite Things / Easy To Love / Baubles, Bangles And Beads |
| Evans | Peace Piece |
| Evans | All The Things You Are |
| Evans | Ornithology (Live) |
| Evans | Medley: "Spartacus" Love Theme / Nardis |
| Evans | Santa Claus Is Coming To Town (Live) |
| Evans | We Will Meet Again (Official Audio) |
| Evans | Hullo Bolinas (Live) |
| Evans | Maxine (Official Audio) |
| Evans | All of You - Bill Evans Transcription (on Marian McPartland Piano Jazz) |
| Evans | Medley: Autumn In New York / How About You? (Live) |
| Evans | Since We Met (Live) |
| Evans | Sugar Plum (Official Visualizer) |
| Evans | All of You - Bill Evans Transcription (on Marian McPartland Piano Jazz) |
| Evans | Come Rain Or Come Shine |
| Evans | Come Rain Or Come Shine |
| Peterson | Yesterdays (Live) |
| Peterson | Take the "A" Train (Live) |
| Peterson | Body and Soul (Live) |
| Peterson | Here's That Rainy Day (Live) |
| Peterson | Mirage (Live) |
| Peterson | Someone to Watch Over Me (Live) |
| Peterson | Perdido (Live) |
| Peterson | Bye Bye Blackbird (Live) |
| Peterson | Who Can I Turn To (Live) |
| Peterson | I Should Care (Live) |
| Peterson | Little Girl Blue |
| Peterson | The more I see you |
| Peterson | (Back Home Again in) Indiana - Oscar Peterson (piano solo) - Montreux 1975 |
| Peterson | At Long Last Love |
| Peterson | Oscar Peterson - Blues Etude |
| Peterson | My Funny Valentine (with Herb Ellis & Ray Brown) (Live) |
| Peterson | Conversation 2 |
| Peterson | Old folks |
| Peterson | Hogtown Blues (Live) |

| | |
|---|---|
| Peterson | Boogie Blues Etude |
| Peterson | When I Fall In Love |
| Peterson | Django |
| Peterson | Watch What Happens |
| Peterson | From The Vault |
| Peterson | Old folks |
| Peterson | Autumn Leaves - Oscar Peterson (1972) |
| Peterson | Greatest piano solo ever |
| Peterson | Body and Soul |
| Peterson | Dancing on the Ceiling (Remastered) |
| McPartland | This Time's The Dream's On Me |
| McPartland | A Fine Romance |
| McPartland | Willow Weep for Me |
| McPartland | Twilight World |
| McPartland | Prelude To A Kiss |
| McPartland | The Duke |
| McPartland | Love You Madly |
| McPartland | Easy Living |
| McPartland | Things Ain't What They Used To be |
| McPartland | Turn Around |
| McPartland | My Funny Valentine |
| McPartland | It's You Or No One |
| McPartland | Marian's Motif |
| McPartland | There'll Be Other Times |
| McPartland | The Single Petal Of A Rose (Live) |
| McPartland | Afterglow @ Monterey Jazz Festival 1975 |
| McPartland | Take The 'A' Train (Live) |
| McPartland | Moonglow |
| McPartland | Billie's Bounce |
| McPartland | Without You |
| McPartland | Mood indigo |
| McPartland | Melancholy Mood |
| McPartland | Close your eyes |
| McPartland | I see your face before me |
| McPartland | Close your eyes |
| McPartland | Star Eyes |
| McPartland | Laura |
| McPartland | Alfie |
| McPartland | Small Talk |
| McPartland | Blue Sophisticate |
| Mehldau | My Heart Stood Still |

| | |
|---|---|
| Mehldau | Roses Blue |
| Mehldau | It's All Right With Me (Live) |
| Mehldau | Memory's Tricks |
| Mehldau | Things Behind the Sun |
| Mehldau | C Tune |
| Mehldau | Waltz Tune |
| Mehldau | From This Moment On |
| Mehldau | River Man |
| Mehldau | Secret Love |
| Mehldau | Unrequited |
| Mehldau | Elegy for William Burroughs and Allen Ginsberg |
| Mehldau | Brad Mehldau - Resignation (solo) |
| Mehldau | I Do |
| Mehldau | Amsterdam |
| Mehldau | Brad Mehldau - Waltz for J. B. |
| Mehldau | Prof Brad Mehldau: The Falcon Will Fly Again - Jazzcampus Basel |
| Mehldau | Jigsaw Falling into Place (Live) |
| Mehldau | Prof Brad Mehldau: John Boy - Jazzcampus Basel |
| Mehldau | Brad Mehldau - Blackbird |
| Mehldau | Am Zauberberg |
| Mehldau | Airport Sadness |
| Mehldau | 29 Palms |
| Mehldau | Meditation I - Lord Watch Over Me (Live) |
| Mehldau | Paris |
| Mehldau | Jigsaw Falling into Place (Live) |
| Mehldau | Perugia |
| Mehldau | Brad Mehldau's Epic Performance of Paranoid Android by Radiohead |
| Mehldau | I Saw Her Standing There |
| Mehldau | Dream Brother (Live) |
| Monk | Dinah (Take 2) |
| Monk | Caravan |
| Monk | Sweet And Lovely (Take 2) |
| Monk | Hackensack |
| Monk | Ask Me Now |
| Monk | North Of the Sunset |
| Monk | I'm Confessin (That I Love you) |
| Monk | I Hadn't Anyone Till You |
| Monk | Everything Happens To Me (Take 3) |
| Monk | Monk's Point |
| Monk | These Foolish Things (Remind Me Of You) |
| Monk | Introspection |

| | |
|---|---|
| Monk | Body and Soul |
| Monk | Just a Gigolo |
| Monk | Thelonious Monk - Reflections |
| Monk | Between the Devil and the Deep Blue Sea |
| Monk | Thelonious Monk ~ North Of The Sunset (1964) |
| Monk | Thelonious Monk - Wee See |
| Monk | Between the Devil and the Deep Blue Sea |
| Monk | Thelonious Monk - I Should Care |
| Monk | Memories of you (take 1) |
| Monk | thelonious monk - don't blame me |
| Monk | Thelonious Monk Piano Solo - Coming On The Hudson |
| Monk | Thelonious Monk Piano Solo - Smoke Gets in Your Eyes |
| Monk | Thelonious Monk Piano Solo - Well, You Needn't |
| Monk | Jazz Portrait Thelonious Monk 1970 - Round Midnight |
| Monk | Thelonious Monk: Off Minor (Paris Piano Solo, 1954) |
| Monk | Thelonious Monk: Evidence (Paris Piano Solo, 1954) |
| Monk | Round Lights by Thelonious Monk from 'Thelonious Alone In San Francisco' |

List of songs from each of the six jazz pianists that 10-second clips were drawn from. Clips that contained only piano solos and improvisation were drawn at approximately one and two-thirds of each song. 10-second clips were then shown to participants during the experimental task.