

Cognitive Psychology

Incidentally Acquiring Pitch-label Associations With a Musical Contingency Learning Task

Claudia Iorio^{1a}, Emmanuel Bigand¹, James R. Schmidt¹¹ Department of Psychology, LEAD-CNRS UMR5022, Université Bourgogne Franche-Comté, Dijon, France

Keywords: pitch identification, contingency learning, music learning, instruction, pitch

<https://doi.org/10.1525/collabra.118806>

Collabra: PsychologyVol. 10, Issue 1, 2024

Incidental learning occurs rather rapidly and effortlessly in a range of different domains, such as early language acquisition, motor learning, and a wide range of more arbitrary laboratory tasks. The present report explores the efficacy of an incidental learning task in the acquisition of pitch-label associations, that is, the ability to identify and name musical notes by ear. In experiment 1, 2 and 3 participants were asked to respond to the target (a note name) while ignoring the cues (either a tone or, in one experiment, a tone with a note position). In a pretest and posttest, we further analyzed their ability to guess the name of the tone in a tone naming task.

We also explored the role of intentionality in acquiring and remembering pitch-label associations, but there were only small suggestive trends for slightly better performance for a group instructed to try to learn the contingencies compared to a purely incidental learning group (i.e., with no instructions about the contingencies), suggesting that learning is at least primarily incidental. Our research opens up new venues for the investigations of incidental learning related to the acquisition of musical features useful to performance (how to play).

Introduction

Although deliberate learning (e.g., formal education) plays an important role in our acquisition of new skills and knowledge, we also learn a considerable amount more passively. *Implicit learning* occurs (a) without the intentional goal to learn and (b) without conscious awareness of what has been learned (Cleeremans et al., 1998; Reber, 1989). These two features of implicit learning are not necessarily perfectly correlated. In the present report, although we do take some awareness measures, we are more interested in *incidental learning*, that is, learning that occurs without the explicit intention to learn (Kerka, 2000), which may or may not lead to unconscious knowledge. Incidental learning occurs rather rapidly and effortlessly in a range of different domains, such as early language acquisition, motor learning, and a wide range of more arbitrary laboratory tasks (e.g., Lewicki, 1986; Nissen & Bullemer, 1987; Oberauer et al., 2015; Reber, 1967; Saffran et al., 1996, 1997, 1999; Turk-Browne et al., 2005).

Perhaps one of the simplest examples of an incidental learning procedure, and a partial inspiration for the present work, is the colour-word contingency learning paradigm (Schmidt et al., 2007; for related learning procedures, see Carlson & Flowers, 1996; Miller, 1987; Mordkoff & Halter-

man, 2008; Musen & Squire, 1993). Contingency learning refers to the ability to detect regularities between events in the environment (e.g., Event B tends to follow Event A, making Event A a predictive cue for Event B; for reviews, see De Houwer & Beckers, 2010; MacLeod, 2019; Schmidt, 2021). In the colour-word contingency learning procedure participants are exposed to regularities between the non-target word stimulus and the target color in which the word is presented. For example, the word “move” might be presented frequently in blue, but rarely in green or red. Although participants are instructed to simply identify the colours (e.g., with a key press) and are not instructed about the associations between words and colours, extremely rapid learning occurs: after few trials, responding is robustly faster and more accurate to trials coherent with the regularity, termed *high contingency* (e.g., “move” in blue), than to trials incoherent with the regularity (e.g., “move” in red), termed *low contingency*.

In similar tasks, many different stimulus dimensions have been used for both the task-irrelevant cue (e.g., shapes, words, nonwords, colors) and task-relevant target (e.g., colors, color words, neutral words, positive/negatively-valenced words; Forrin & MacLeod, 2017; Levin & Tzelgov, 2016; Schmidt & De Houwer, 2012a, 2012b, 2012c). Learning is always very rapid and the pattern of re-

a Correspondence: claudia.iorio@cnrs.fr

sults is always the same, that is, faster and more accurate responding to high relative to low contingency trials. Similarly fast learning is observed in other incidental learning procedures such as sequence learning (Bianco et al., 2020; Nissen & Bullemer, 1987; Turk-Browne et al., 2005; Woods & McDermott, 2018), artificial-grammar learning (Reber, 1967; for a review, see Pothos, 2007), the Hebb digits task (McKelvie, 1987; Oberauer et al., 2015; Vachon et al., 2018), and hidden covariation detection (Lewicki, 1985, 1986; Lewicki et al., 1992).

The present report explores the efficacy of an incidental learning task in the acquisition of pitch-label associations, that is, the ability to identify and name musical notes by ear. Although incidental learning has been robustly observed with a wide range of different stimulus materials, learning to name musical pitches by ear is a particularly interesting case. As discussed below, in the music cognition and musicology literatures this particular skill is considered to be especially difficult to master, perhaps even impossible for most adults. This might suggest a particular (and surprising) boundary condition on incidental learning. On the other hand, if learning to identify pitches by ear is learnable incidentally, the present work might suggest future avenues for aiding novice musicians in acquiring this skill.

Incidental Learning in Music

Formal instruction and deliberate learning obviously are fundamental for understanding musical theory and becoming an expert musician. However, most people, even non-musicians, possess some music competences that they gained from mere exposure (Bigand & Poulin-Charronnat, 2006; Rohrmeier & Rebuschat, 2012). For instance, we can all easily recognize and correctly reproduce (e.g., by humming) a familiar melody without having explicit knowledge of the music grammar. The incidental (and implicit) learning of musical material has been already investigated in prior work, such as the acquisition of sequence information linked to melody (Saffran et al., 1999, 2000; Tillmann & Poulin-Charronnat, 2010), timbre (Bigand et al., 1998), harmony (Bly et al., 2009; Loui et al., 2009; Rohrmeier & Cross, 2009), and rhythm (Brandon et al., 2012; Salidis, 2001; Schultz et al., 2013; Tillmann et al., 2011). Much as with more arbitrary materials, such as in colour-word contingency learning, artificial grammar learning, sequence learning, or hidden covariation detection tasks, regularities in musical materials are also learned rapidly and robustly.

Recently, in a series of studies (Iorio et al., 2023; Schmidt et al., 2023), we applied a similar logic to the more ecological case of acquiring sight-reading skills. Using a musical contingency learning procedure, participants were asked to identify note names (the relevant stimulus or target) while ignoring note positions (the irrelevant stimulus or cue). Critically, each note position was presented much more frequently with the congruent note name (e.g., “do” was written inside of the note position for “do” much more often than incongruent note names). Although the participants were not informed about or instructed to pay attention to these contingencies, nonmusicians learned note name/note position associations and they were able to correctly use

their knowledge in a note naming test. Again, learning was very fast. The entire experiment lasted about 20 minutes and robust learning was already observed within this period.

One of the reasons why incidental learning appears so quickly is the large numbers of trials that participants experience in a very short time. That is, participants gain substantial practice with novel stimuli rapidly. For instance, in some of our music learning studies mentioned above, participants saw 336 trials in roughly 15 minutes. As such, this type of learning procedure allows for rapid automatization. We also saw this in our performance measures. For example, nonmusicians participants responded robustly faster to congruent than to incongruent trials during the learning phase. This suggests that participants have not only learned the meanings of the note positions, but that seeing a note position provokes a very rapid retrieval of the corresponding note name.

Pitch Identification: A Special Case?

Surprisingly, the same types of incidental learning tasks have not been used to explore whether participants are able to learn to identify pitches by ear and to internalize their pitch identities. This is a particularly interesting question, both theoretically and practically, because the ability to identify and name pitches by ear is considered to be so difficult that it may be unlearnable by most. In particular, *absolute pitch* (AP) is the ability to name a pitch by ear (for reviews, see Deutsch, 2013; Levitin, 2007). For instance, an AP possessor can hear a random note played on an instrument without any initial context (e.g., hearing an initial note of known pitch) and the AP possessor would be able to correctly name the note (e.g., “mi”), usually very rapidly and with little effort.

In the current report, we do not study AP possessors or the ability to acquire true AP. However, the difficulty of acquiring AP suggests that there might be something fundamentally difficult about learning pitch-label associations. Although the criteria for determining what “counts” as AP varies rather unsystematically in the literature, AP possessors identify pitches quite accurately (Levitin & Rogers, 2005; Miyazaki, 1988). Their pitch identification is not always perfect, but their errors tend to be very close to the correct response (e.g., ± 1 semitone). Further, their identification of pitches by ear is very automatic, with response times generally between 1.5 and 3 seconds according to some reports (Bermudez & Zatorre, 2009; Miyazaki, 1990; Takeuchi & Hulse, 1993; Van Hedger et al., 2019; Wong, Lui, et al., 2020), or even as rapid as 600 ms according to others (Refaat, 2014).

AP ability is rare, only present in a small percentage of the population (Miyazaki et al., 2012; Takeuchi & Hulse, 1993; Ward, 1999), even among skilled musicians. Some authors propose that there is a strong genetic component (Athos et al., 2007), supported by twin studies (Theusch & Gitschier, 2011), early acquisition (Deutsch, 2013), and unique structured brain circuitry (Bermudez & Zatorre, 2009; Loui et al., 2009; Schulze et al., 2009). Others have suggested that there is a critical period, with AP rarely ob-

served for those starting music training after 4 or 5 years old (Crozier, 1997; Deutsch, 2013; Deutsch et al., 2006; Miyazaki & Ogawa, 2006), similar to the way that the ability to make phonemic distinctions that do not exist in the native language rapidly reduces after a critical period in early language learning (e.g., Werker & Tees, 1984).

The above-mentioned research on AP suggests an interesting question: Are the associations between auditory pitches and their corresponding pitch names fundamentally impossible (or very difficult) to learn (e.g., for adult AP non-possessors)? For instance, is there something “special” about musical notes that makes it almost impossible to learn how to associate a note name with them? If the response to this type of question is affirmative, then this would seem rather surprising from the lens of research on incidental learning. As long as the regularity to learn is relatively simple (e.g., as is the case when learning simple pairings between auditory pitches and note names), learning tends to be rather rapid and robust and does not seem to have a strong modality sensitivity. Whether we are asking participants to learn an artificial grammar from an auditory artificial speech stream, to learn pairings between words and colours, between nonwords and emotional stimuli, or a wide range of other types of stimulus pairings, participants generally have no difficulty learning such regularities. Why would note pitches (or the relationships between note pitches and their corresponding names) be any different? To explore this question in the present work we developed an incidental learning task that is structurally similar to other types of incidental learning tasks with nonmusical (or other types of musical) materials. We then ask whether the same type of rapid and effortless learning is observed in a pitch-label learning task or whether there is evidence for a fundamental difficulty in learning this specific type of stimulus pairing.

Pitch Identification: Not So Special?

Some research does suggest that the ability to name pitches by ear is easier than typically assumed, at least at a more implicit level. To better appreciate what has been observed in research on “implicit AP” (Deutsch, 2013; Levitin, 2007; Schellenberg & Trehub, 2003), it is first relevant to make a distinction between absolute and relative pitch. *Relative pitch (RP)* is the ability to identify and name pitches after receiving an external reference. Concretely, if we play a context note and inform the RP possessor of its identity (e.g., “fa”) and then play a different note (e.g., the pitch for “la”), then the RP possessor should be able to identify the second note. RP possessors achieve this with a comparison-based strategy (Levitin, 1994; Levitin & Rogers, 2005; Takeuchi & Hulse, 1993). They have therefore not learned the name of each auditory pitch but can rather “calculate” the correct pitch name for a note via a comparison with the context note of known pitch. Because of this, pitch naming is much less automatic and rapid in RP possessors.

Work on what has sometimes been termed *implicit AP* indicates that most people, even though categorized as “AP non-possessors”, are able to succeed at tasks that should be impossible without AP. As an illustration, AP non-pos-

sessors are able to judge whether a familiar piece of music is played in the correct key (Miyazaki & Rakowski, 2002), which should not be possible with RP alone. That is, if a piece of music is simply transposed, for instance, from C Major to B Major, all the intervals between notes remain identical. As such, detecting that one version is correct (e.g., C Major) and that the other is incorrect (B Major) necessarily requires detecting pitches absolutely. Similarly, AP non-possessors can correctly reproduce familiar melodies (e.g., by humming) with a reasonable degree of accuracy (Levitin, 1994). Again, this should only be possible with AP. RP would be insufficient to produce the pitches absolutely, and many of the participants studied in this type of research possessed neither (i.e., non-musicians without AP or RP). These results might be taken to suggest that there is something fundamentally wrong with the idea that pitch-label associations are nearly impossible to learn. Alternatively, it might be proposed that pitch-label associations are learnable implicitly but not explicitly.

However, some other recent studies (Van Hedger et al., 2019; Wong, Lui, et al., 2020; Wong, Ngan, et al., 2020) have hinted that some adults might be able learn AP, even at an explicit level. They demonstrated that after explicit, extended, and effortful training, adults (musicians and non-musicians) were able to improve their speed and accuracy in pitch identification tasks. Some showed performance at posttest similar to true AP possessors. This work is not without its critics, however. For instance, participants with accurate posttest scores often had pretest scores that were already reasonably good, thus demonstrating only moderate improvements. It is not necessarily controversial to suggest that pitch identification can be improved, but the general consensus seems to be that such improvements are likely to be minimal and that it is implausible to think that someone without any pitch identification abilities at all could learn to easily and rapidly identify pitches beyond some of the stricter criteria.

In any case, our goals are notably different than the pre-existing research discussed above. Past work has used extended and explicit training to determine whether some participants are able to achieve AP-level performance after training and how much of an improvement is possible. The current work does not aim to address such questions. Rather, our goal is to determine whether rapid and robust acquisition and improvement of pitch-label associations is possible at all in an incidental learning task. In particular, much of the research discussed above would suggest that pitch identification is uniquely difficult. As such, the procedures that work quite rapidly for learning other types of associations (i.e., incidental learning procedures) may not be nearly as effective when trying to acquire pitch-label associations. Our basic postulate, however, is that this notion is likely false, and that rapid learning and improvement of pitch-label associations should be possible. We also note in advance that the learning we observe may or may not be comparable to true AP perception, a point to which we will return in the General Discussion.

Table 1. Contingency manipulation

Note Name	Tone						
	fa	sol	la	si	do	ré	mi
fa	18	1	1	1	1	1	1
sol	1	18	1	1	1	1	1
la	1	1	18	1	1	1	1
si	1	1	1	18	1	1	1
do	1	1	1	1	18	1	1
ré	1	1	1	1	1	18	1
mi	1	1	1	1	1	1	18

Note: In the table is represented the contingency proportion between high contingency trials (presented 80% or 18 times) and low contingency trials (presented 20% or 1 time each). For instance, the tone “fa” is presented much more often with the note name “fa” (high contingency trials) than with the other note names (low contingency trials).

Current Work: The Music Contingency Learning Procedure

The goal of the present work is to explore the learnability of pitch-label associations in an incidental learning task. We note in advance that we take a simplified approach to studying early learning. For instance, artificial grammar learning studies do not have participants learn the entire grammar of an entire language, but rather researchers create a limited set of grammar rules with a small number of letters to assess the degree to which grammar rules are learnable incidentally (for instance, see Saffran et al., 1999). Similarly, in the present work, we do not train participants with all the semitones from multiple octaves and timbres (i.e., instruments) and test them with strict tests of AP (e.g., Van Hedger et al., 2019; Wong, Lui, et al., 2020; Wong, Ngan, et al., 2020). Instead, we train participants with a smaller number of stimuli in a task highly similar to other incidental learning tasks (e.g., Iorio et al., 2023; Schmidt et al., 2007) to see whether pitch-name associations pose a particular difficulty for participants and to what degree participants can automatize pitch-label associations.

Auditory musical Stroop procedures are one way to easily assess the automaticity of pitch processing (Akiva-Kabiri & Henik, 2012; Hamers & Lambert, 1972; Leboe & Mondor, 2007). Generally, in these procedures participants are asked to respond to a relevant stimulus while ignoring an irrelevant stimulus that is either congruent (e.g., the word “high” presented in a high-pitched voice) or incongruent (e.g., the word “high” presented in a low-pitched voice), analogous to colour-word Stroop tasks (see MacLeod, 1991, for a review) or sight-reading music Stroop tasks (see Grégoire et al., 2013). Faster RTs for congruent trials compared to incongruent trials indicates that pitch processing is automatic. That is, although participants are asked to respond to the words, they cannot avoid processing the pitch, resulting in slower RTs when the association between the stimuli is incongruent. In one experiment, Akiva-Kabiri & Henik (2012) compared performance in a tone naming task and note naming task between AP possessors and non-possessors. In the tone naming task, participants were asked

to respond to the tone while ignoring the note name. In the note naming task, participants were asked to do the reverse (i.e., to respond to the note name while ignoring the tone). They found a congruency effect for AP possessors only in the note naming task and a congruency effect in the tone naming task for non-AP possessors, suggesting that only AP possessors are automatically biased by pitches when identifying note names.

Of course, this work compared those with pre-existing pitch identification skills (i.e., AP possessors) with those that do not already possess such skills (i.e., AP non-possessors). Our goal, in contrast, is to study learning of pitch detection abilities. It was our hypothesis that participants can not only be trained to improve accuracy in pitch identification but will also show evidence of automaticity in performance measures. In the following studies we use an auditory adaptation of the above-mentioned musical contingency learning task (Iorio et al., 2023) to measure the automaticity of pitch processing in nonmusicians and musicians. Analogous to the manipulation we used in the sight-reading learning procedure, participants heard a tone (cue) and then they were asked to respond to the note name (target) that appeared in the center of the screen. The note name was presented much more often with its congruent tone (e.g., the name “do” with the tone “do”) than with any other of the incongruent tones (e.g., the name “mi” with the tone “do”), as illustrated in [Table 1](#).

Our key hypothesis is that participants will be able to learn (or improve) their pitch identification abilities. This should be reflected both in an increase in explicit identification of note pitches after training, and more automatic effects on performance (i.e., faster responses to high-contingency congruent trials relative to low-contingency incongruent trials).

Experiment 1

In Experiment 1, we wanted to investigate whether non-musicians were able to incidentally learn pitch-label associations. Nonmusicians are an interesting group to study, because normally they will have little or no practice with pitch identification. They are thus a naïve control group, and similar also to beginner musicians. For this purpose, we

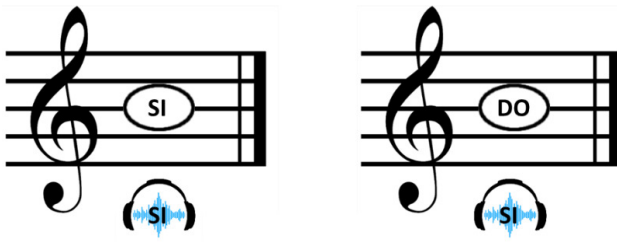


Figure 1. Example of high and low trials

Note. An example of high/congruent trial on the left in which the tone, the note position and the note name matched. On the right there is an example of low/incongruent trial where the tone and the note position matched between themselves, but not with the note name.

used a modified version of the musical contingency learning procedure from our previous studies (Iorio et al., 2023; Schmidt et al., 2023), as discussed in the Introduction. Because previous research suggested that a combined presentation of both note positions and tones can benefit the acquisition of musical skills such as sight-reading (Mishra, 2014), one might posit that learning to identify pitches by ear would also be improved by presenting musical notation as a supplementary visual cue. We therefore compared two groups that were exposed to different cue-target associations. In the tone-cue group, only tones were used as cues (i.e., the only visual stimulus was the target note name). In the multiple-cues group, however, both note positions and tones were used as cues. Specifically, participants were presented with a musical staff. A note was presented in one of the positions of the music staff at the same time as the tone. The note position and tone always matched. As in the tone-cue group, the tone (and note position) was predictive of the target note name, the latter of which was presented inside of the note position.

Our primary hypothesis is that both groups of nonmusicians will incidentally learn the pitch-label associations. We hypothesize that participants will learn the associations quickly, showing both improved accuracy in explicit pitch identification during the test phase and automatic effects on performance during the learning phase. In particular, we anticipate faster responses to congruent (high-contingency) trials than to incongruent (low-contingency) trials (see Figure 1 for an example of high- and low- contingency trials). Concerning the group factor, we considered two contrasting hypotheses. First, we might expect larger learning effects in the multiple-cues group compared to the tone-cue group. The combination of the note positions along with the tones might reinforce learning of the tone-label associations. On the other hand, another possibility is that adding in a second cue actually impairs learning about the tone-label associations. This might result if there is overshadowing (Pavlov, 1927). *Overshadowing* is the observation that the learning of one regularity is impaired by the simultaneous learning of another regularity. Specifically, the presence of associations between note positions and note names might impair the learning of associations between tones and note names, and this because participants learn the regularities between the note positions and note names *instead of* the associations between pitches and note

names (for more discussion of theories of overshadowing, see the General Discussion). This group factor was largely exploratory, as we did not have strong a priori prediction for either of the two contrasting hypotheses mentioned above.

Method

Participants

119 participants, recruited online on [Prolific.co](https://prolific.co), were randomly assigned to one of the two experimental conditions described below (59 participants in the multiple-cues group and 60 in the tone-cue group) and received monetary compensation (3.80 £) for their participation. Our inclusion criteria, mentioned in the recruitment advertisement, were being able to understand French, being between 18-30 years old, not being a musician, and not being able to read musical notation. 16 participants reported having absolute pitch. Precisely, 15,25% participants (9 of 59) in the multiple-cues group and 11,66% participants (7 of 60) in the tone-cue group answered yes to the subjective awareness question regarding absolute pitch. Overall, their performance on the pretest (in which they were asked to guess the name of the tones) were not significantly higher than the performance of the remaining 103 participants that did not claim to have perfect pitch: $t(117) = .044, p = .946, d = .012, BF_{10} = .271, M_{\text{absolute pitch participants}} = 16.7\%, SD = 9.52, \text{highest score} = 33.33\%; M_{\text{remaining participants}} = 16.8\%, SD = 14.0, \text{highest score} = 85.71\%$. Therefore, we did not exclude these participants from the analysis. However, although three participants in the tone-cue group declared that they did not have absolute pitch, their performance in the pretest were between 60% and 100%, similar to AP possessors' performance reported in the literature (Levitin & Rogers, 2005; Miyazaki, 1988). For this reason, these participants were excluded from the following analysis.

Ethical review and approval were not required for the study on human participants in accordance with the local legislation and institutional requirements. All participants accepted a written consent before beginning the study. All the procedures were conducted in accordance with the Declaration of Helsinki. Participants' anonymization was guaranteed.

Apparatus, Design, and Procedure

The experiment was programmed and run with PsychoToolkit, a web-based software that allows reliable RTs as shown from previous research (Stoet, 2010, 2016), also with musical stimuli (Armitage & Eerola, 2020). The auditory stimuli were pure sinewaves that were created using Audacity software with the lowest pitch being the "fa" (or "F") note at the frequency of 349.228 Hz and the highest pitch being the "mi" (or "E") note at the frequency of 659.255 Hz. The "la" (or "A") pitch was thus tuned to the standard tuning at the frequency of 440 Hz. To ensure that headphones or speakers were correctly working during the task, participants completed a sound check before starting the experiment.

During the main parts of the experiment, participants responded with the Z-I keys on a standard AZERTY keyboard. However, because the experiment was online and it involved participants from different countries (though AZERTY is standard in French-speaking countries), an instruction referring to the type of keyboard needed in the study was added in the recruitment advertisement. The keys Z, E, R, T, Y, U, and I were labelled according to the sequence of the musical scale from the lower to upper position (i.e., *fa, sol, la, si, do, ré, and mi*, respectively, referring to the French note names). The “O” and “N” keys were additionally used to answer “Oui” (Yes) or “Non” (No) to the subjective awareness question, and the spacebar was used to begin each phase from the instruction screens.

Before starting the experiment, we collected a subjective measure for AP in which participants were asked whether they were able to name a tone without previously listening to a reference note, translated from French:

“Do you have perfect pitch, which means that you can name one or more tones when listening without first having to hear an identified note serving as a reference?”

This question was primarily used for screening purposes, along with the pretest scores, as described above in the Participants section.

The experiment started with two practice phases, in which participants practiced and automatized the note name-to-key assignments. During these phases participants were presented only with the note names. The trial started with a fixation cross (“+”) in the center of the screen for 500 ms, followed by a blank screen for 250 ms. A French note name (*fa, sol, la, si, do, ré, or mi*) was then presented in the center of the screen until response (no time limit). An on-screen key reminder (Z, E, R, T, Y, U, I) was added throughout the first practice phase to help participants to learn the note name-to-key assignments. Following correct responses, the next trial began immediately. Following incorrect responses, the note name changed color to red and stayed on the screen until the participant pressed the correct key. The second practice phase was identical in all respects, except that the on-screen key reminder was removed and participants were encouraged to try to respond from memory. There were 70 trials in each practice phase (140 trials in total).

Before beginning a test phase, participants were randomly assigned to one of the two groups (multiple-cues and tone-cue groups). While in the multiple-cues group the target was preceded by both a note position and a tone (predictive cues), in the tone-cue group the note name was only preceded by a tone. The procedure was otherwise identical for the two groups, with exceptions noted below. A pretest phase, which measures the ability of the participant to discriminate (e.g., better-than-chance guessing) between experienced and unexperienced events (Cheesman & Merikle, 1984), followed the practice phases. The pretest (42 trials in total) allowed us to assess the ability of participants to identify tones (and note-positions in the multiple-cues group) prior to learning. Specifically, we were in-

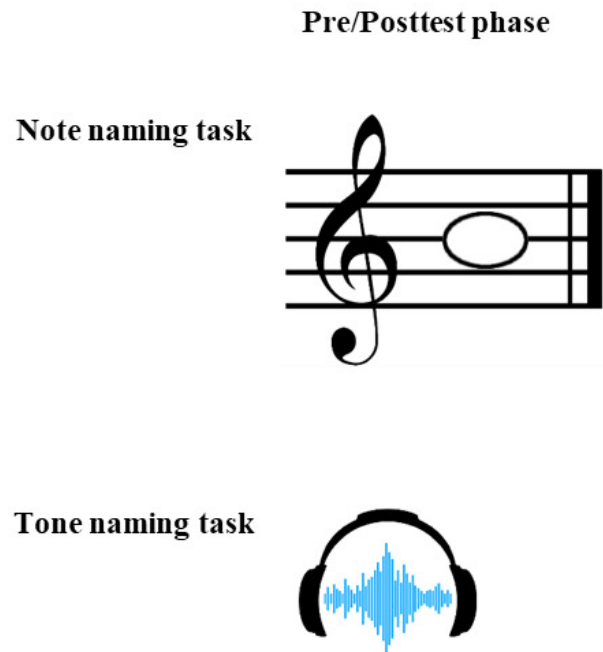


Figure 2. Schematic description of the pre/posttest phases

Note. Both groups were presented with the tone naming task, in which they had to guess the name of the tone. The note-position naming task was presented only in the multiple-cues group.

terested in knowing whether our participants were able to recognize and name the tones (and note-positions) used as our predictive cues before starting the learning phase. As previously mentioned, Experiment 1 was conducted with nonmusicians as a sort of pure control group, who should normally have no pitch identification (or sight-reading) skills in the absence of music training, but the pretests allowed us to both (a) screen for undisclosed pre-existing knowledge and (b) to establish a control for pre/post improvement scores. While we used both note positions and tones as predictive cues for the multiple-cues group, only the tones preceded the note name in the tone-cue group. Therefore, both groups were presented with the tone naming task (Figure 2), in which they had to guess the name of the tone (no limit time; 21 trials). The note-position naming task (Figure 2), in which a music staff appeared in the center of the screen for 500 ms, then a note position appeared on the staff until participants responded (no limit time; 21 trials), was presented only in the multiple-cues group.

Immediately after the pretest phase, participants started the learning phase that differed between the groups as shown in Figure 3. The multiple-cues group was presented, on each trial, with a musical staff that appeared on the screen for 500 ms. The note was then added to the staff and the tone started playing for 250 ms. The note name was then written inside the note and participants had 3000 ms to respond. After the note name was presented, the tone continued playing for another 500 ms (750 ms total) or until a response was made. Following correct responses, the next trial began immediately. If participants responded incor-

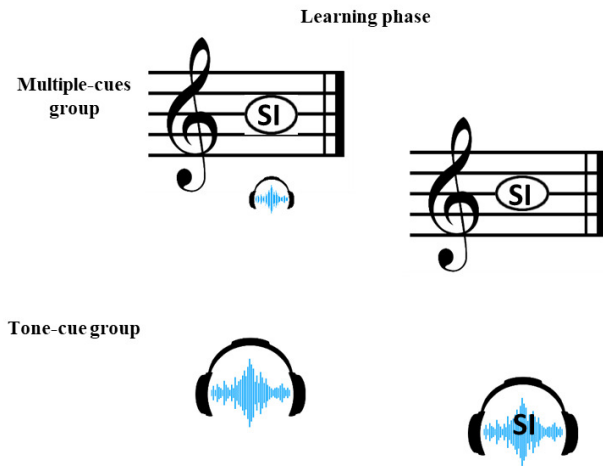


Figure 3. Schematic description on the learning phase

Note. On the top, an example of how the learning phase looked like for the multiple-cues group. On the bottom the learning phase for the tone-cue group.

rectly or failed to respond in 3000 ms, the note name was replaced with “XXX” in red for 500 ms before the beginning of the next trial. Globally, the same structure was also used for the tone-cue group, with a few exceptions: only the tone was presented as predictive cue (instead of both tone and note position), no musical staff was presented on the screen, and a fixation cross was presented in the center of the screen from the tone onset until it was replaced by the note name. In total, there were 420 trials in the learning phase, randomly ordered (without replacement), and a contingency manipulation of 90% (Schmidt et al., 2023) congruent pairings (e.g., the tone “fa” for the note name “fa”; high-contingency trials) and 10% incongruent trials (e.g., the tone “fa” for the note name “do”; low-contingency trials), as illustrated in Table 2. The congruency (or contingency learning) effect was measured as the difference in response times or error rates between low- and high-contingency trials.

Following the main learning phase, contingency awareness was collected to assess whether participants noticed the regularities during the learning phase. In particular, participants were assessed for *subjective awareness* (Cheesman & Merikle, 1984). For this, they responded to

an on-screen instruction, where it was asked if they noticed that some pairings (high-contingency trials) were presented more often than others (low-contingency trials).

Participants could respond “yes” or “no” with a key press. This screen read (translated from French):

“During the third part of this experiment, note names were presented with a tone (or with a tone and a note position for the multiple-cues group). Each tone was presented more frequently with one note name than the others. That is to say, one tone was frequently presented with “do,” another frequently with “re,” etc. Did you notice these regularities?”

Directly after, the posttest phase started and it was exactly the same as the pretest phase. This allowed us to compare participants’ performance before and after the learning process. The instructions for these phases were (translated from French):

“Now, the task is similar, except that you will only hear a tone. Try to guess the name of the tone by pressing the appropriate key on the keyboard.”

A slightly different instruction was presented to the multiple-cues group (translated from French):

“Now, the task is similar, except that you will only see a note and hear a tone. Try to guess the name of the note and the tone by pressing the appropriate key on the keyboard.”

Data Analysis

We conducted analyses on the learning and the test phases. For the learning phases, we conducted a repeated measures ANOVA on correct RTs and error rates to assess the overall main effects of contingency, group, and the interaction between them. Trials in which participants failed to respond in 3000 ms (i.e., before the deadline) were eliminated (on average on all the 119 participants, 12.54% of the trials were eliminated). For the test phases, we analyzed accuracy rates to assess whether participants responded above chance (the chance guessing rate was 1/7 or approximately 14.3%) and response times. All analyses were evaluated at the $\alpha = .05$ level of significance. Additionally,

Table 2. Contingency manipulation

Note Name	Tones						
	fa	sol	la	si	do	ré	mi
fa	54	1	1	1	1	1	1
sol	1	54	1	1	1	1	1
la	1	1	54	1	1	1	1
si	1	1	1	54	1	1	1
do	1	1	1	1	54	1	1
ré	1	1	1	1	1	54	1
mi	1	1	1	1	1	1	54

Note: In the table is represented the contingency proportion between high-contingency trials (presented 90% or 54 times) and low-contingency trials (presented 10% or 1 time each). For instance, the tone “fa” is presented much more often with the note name “fa” (high-contingency trials) than with the other note names (low-contingency trials).

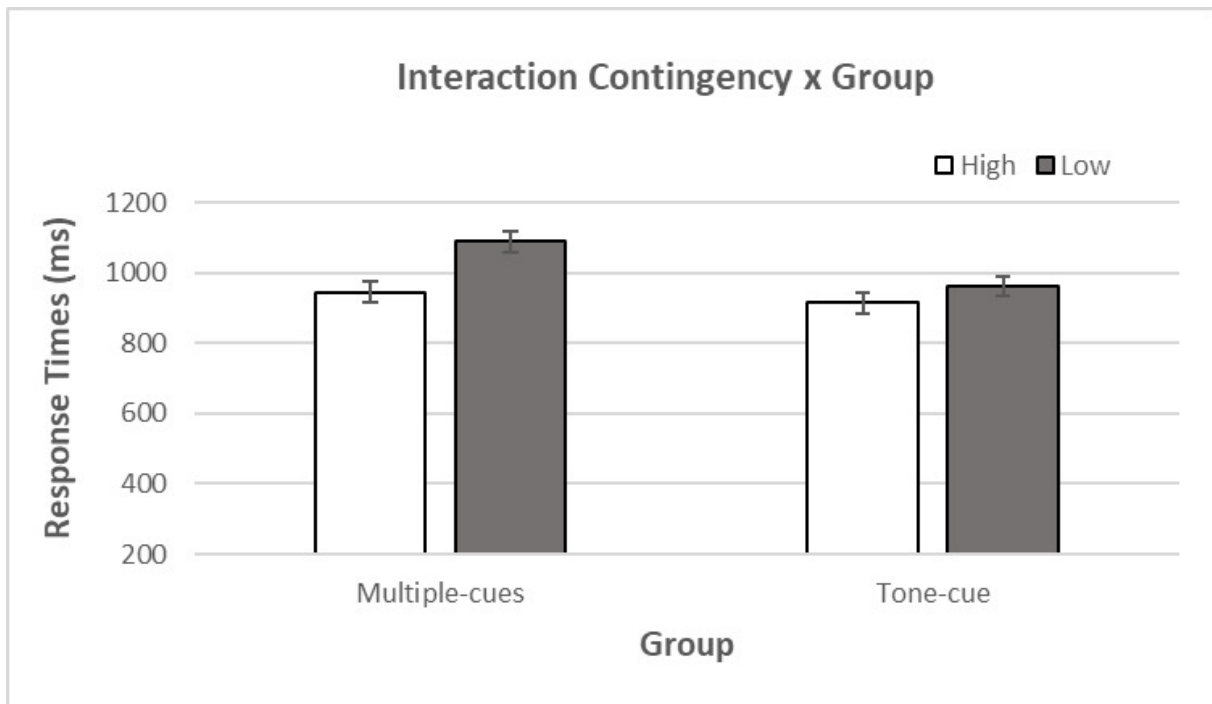


Figure 4. Interaction between contingency effect and Group in Experiment 1

Note. Interaction between Contingency (High and Low) and Group (multiple-cues vs. tone-cue), standard error bars are shown in the figure.

we consistently reported the Bayes factor, computed using JASP software (JASP Team, 2019). We used the standard noninformative Cauchy prior with a default width of 0.707. We report the Bayes factor BF_{10} , with values between 3 and 10 supporting moderately strong evidence for the alternative hypothesis (H1; Doorn et al., 2021). The data set is available via the following link: <https://osf.io/xjdt4/>.

Results

Response Times

Response time results are presented in Figure 4. A repeated measures ANOVA for RTs with the factors Contingency (high vs. low) and Group (multiple-cues group vs. tone-cue group) indicated a significant main effect of Contingency, $F(1,114) = 74.0$, $p < .001$, $\eta^2 = .394$, $BF_{10} > 100$, showing faster responses for high-contingency trials ($M = 935\text{ms}$, $SD = 231$) than for low-contingency trials ($M = 1027\text{ms}$, $SD = 224$). The main effect of Group was not significant, $F(1,114) = 3.33$, $p = .071$, $\eta^2 = .028$, $BF_{10} = 1.18$. The interaction between Contingency and Group was significant, $F(1,114) = 25.2$, $p < .001$, $\eta^2 = .181$, $BF_{10} > 100$, indicating a greater difference between high and low contingency trials in the multiple-cues group ($M_{\text{high trials}} = 946\text{ms}$, $SD = 257$; $M_{\text{low trials}} = 1089\text{ms}$, $SD = 232$) compared to the tone-cue group ($M_{\text{high trials}} = 925\text{ms}$, $SD = 202$; $M_{\text{low trials}} = 962\text{ms}$, $SD = 199$). The contingency effect was significant for both the multiple-cues group, $M_{\text{low-high}} = 143$, $SD = 139$; $t(58) = 7.92$, $p < .001$, $d = 1.03$, $BF_{10} > 100$, and the tone-cue group, $M_{\text{low-high}} = 37.6$, $SD = 78.0$; $t(56) = 3.64$, $p < .001$, $d = .482$, $BF_{10} > 100$.

Error Rates

The repeated measures ANOVA for errors with the factors Contingency (high vs. low) and Group (multiple-cues group vs. tone-cue group) revealed a significant main effect of Contingency, $F(1,114) = 33.5$, $p < .001$, $\eta^2 = .227$, $BF_{10} > 100$, and a non-significant main effect of Group, $F(1,114) = 3.38$, $p = .068$, $\eta^2 = .029$, $BF_{10} = 1.06$. The interaction between Contingency and Group was also significant as shown in Figure 5, $F(1,114) = 12.5$, $p < .001$, $\eta^2 = .099$, $BF_{10} = 43.84$ (multiple-cue group: $M_{\text{high}} = 12.3\%$, $SD = 8.24\%$, $M_{\text{low}} = 20.4\%$, $SD = 13.2\%$; tone-cue group: $M_{\text{high}} = 12.0\%$, $SD = 8.23\%$, $M_{\text{low}} = 13.9\%$, $SD = 13.4\%$). The contingency effect was significant in the multiple-cues group, $M_{\text{low-high}} = 8.14\%$, $SD = 10.3\%$; $t(58) = 6.05$, $p < .001$, $d = .788$, $BF_{10} > 100$, and not significant in the tone-cue group $M_{\text{low-high}} = 1.96\%$, $SD = 5.16\%$; $t(56) = 1.78$, $p = .081$, $d = .236$, $BF_{10} = .631$.

Pre/Posttest Phases

Here we report subjective awareness for the tone naming task. In the multiple-cues group, 67,80% (40 of 59) of the participants noticed the contingencies between note names and tones. In the tone-cue group, this percentage was 54,39% (31 of 57). The subjective awareness question concerning note positions was only posed to the multiple-cues group (i.e., as the tone-cue group did not see note positions). In the multiple-cues group, 60,40% (38 of 59) participants became aware of the contingencies between note names and note positions.

The t -tests for pretest and posttest accuracy, as shown in Figure 6, showed that in the note position naming task,

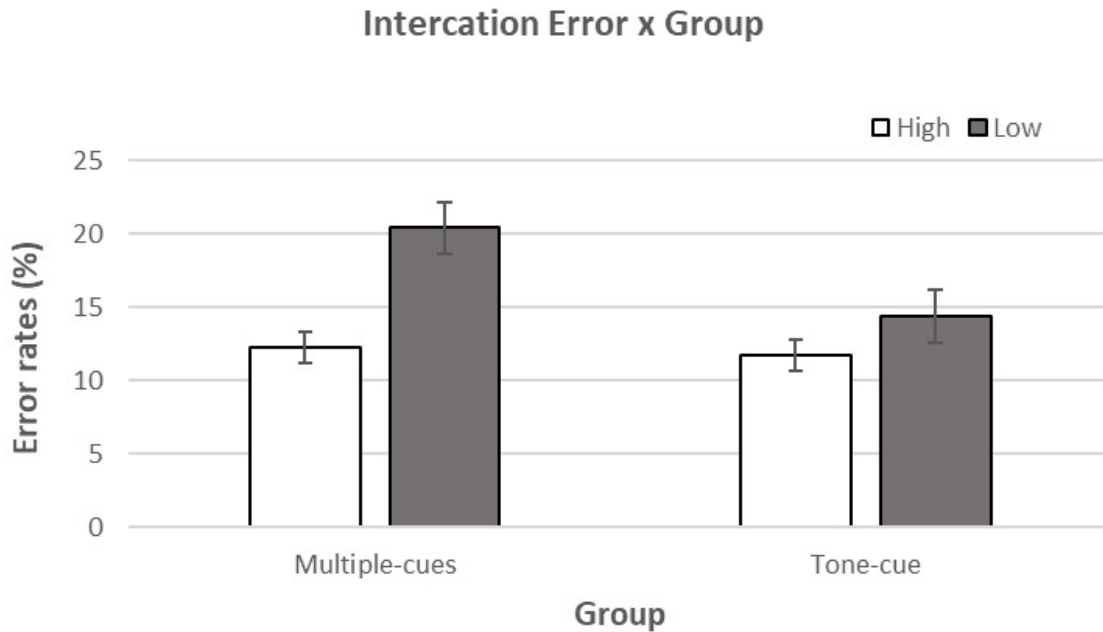


Figure 5. Interaction between Error effect and Group in Experiment 1

Note. Interaction between Error (High and Low) and Group (multiple-cues vs. tone-cue), standard error bars are shown in the figure.

the multiple-cues group performed well above chance (i.e., 14.3%) in both the pretest, $t(58) = 5.95$, $p < .001$, $d = .774$, $BF_{10} > 100$, $M = 37.6\%$, $SD = 30.1\%$, and posttest, $t(58) = 8.32$, $p < .001$, $d = 1.08$, $BF_{10} > 100$, $M = 49.8\%$, $SD = 32.8\%$. Further, performance was significantly improved in the posttest compared to pretest, $t(58) = 3.46$, $p = .001$, $d = .450$, $BF_{10} = 26.5$.

We then ran an ANOVA with the factors Test (pre vs. post) and Group (multiple-cues vs. tone-cue) on the tone naming accuracy rates. The results showed a significant main effect of Test, $F(1,114) = 15.49$, $p < .001$, $\eta^2 = .120$, $BF_{10} > 100$, indicating higher accuracy in naming tones in posttest ($M = 21.1\%$, $SD = 15.0$) relative to pretest ($M = 15.4\%$, $SD = 10.1$). There was also a weak significant main effect for Group, $F(1,114) = 4.30$, $p = .040$, $\eta^2 = .036$, $BF_{10} = 1.04$, indicating higher overall accuracy in the tone-cue group than in the multiple-cues group. More importantly, there was a significant interaction between Test and Group, $F(1,114) = 6.60$, $p = .011$, $\eta^2 = .055$, $BF_{10} = 5.69$, indicating larger improvements in accuracy on posttest in the tone-cue group ($M = 25\%$, $SD = 16.3$) relative to the multiple-cues group ($M = 17.4\%$, $SD = 12.7$), as shown in Figure 6. The multiple-cues group did not perform significantly above chance in the pretest, $t(58) = .824$, $p = .413$, $d = .107$, $BF_{10} = .197$, $M = 15.4\%$, $SD = 10.4\%$, and in the posttest, $t(58) =$

1.898 , $p = .063$, $d = .247$, $BF_{10} = .760$, $M = 17.4\%$, $SD = 12.7\%$. The tone-cue group was also not significantly above chance in the pretest, $t(56) = .814$, $p = .419$, $d = .108$, $BF_{10} = .198$, $M = 15.4\%$, $SD = 9.94\%$, but was significantly above chance in the posttest, $t(56) = 4.94$, $p > .001$, $d = .654$, $BF_{10} > 100$, $M = 25.0\%$, $SD = 16.33\%$. More importantly, the data showed a significant improvement between the pretest and posttest, $t(56) = 4.33$, $p > .001$, $d = .574$, $BF_{10} > 100$, for the latter group.¹

Furthermore, the ANOVA with the factors Test (pre vs. post) and Group (multiple-cues vs. tone-cue) on the RTs in the tone naming task, showed a main significant effect of Test, $F(1,114) = 8.435$, $p < .004$, $\eta^2 = .069$, $BF_{10} = 7.049$, indicating that participants were overall slower in naming tones in the pretest ($M = 2296$ ms, $SD = 1795$) compared to the posttest ($M = 1796$ ms, $SD = 1603$). Both the main effect for Group, $F(1,114) = .386$, $p = .536$, $\eta^2 = .003$, $BF_{10} = .244$, and the interaction between Test and Group, $F(1,114) = .262$, $p = .610$, $\eta^2 = .002$, $BF_{10} = .227$, were not significant.

Overall, while a significant improvement was found between the pretest and the posttest rates in the tone-cue group, the same effect was not observed in the multiple-cues group, potentially indicating an overshadowing effect.

¹ Given some violations of the normality assumptions of the ANOVAs, we also ran a non-parametric repeated measures ANOVA (Friedman test) on the tone naming accuracy (as suggested by an anonymous reviewer). Consistent with the above ANOVA results, no improvements between the pretest and posttest were observed for the multiple-cues group, $\chi^2 = .018$, $p = .893$. In contrast, a significant improvement between pretest and posttest was observed for the tone-cue group, $\chi^2 = 8.00$, $p = .005$.

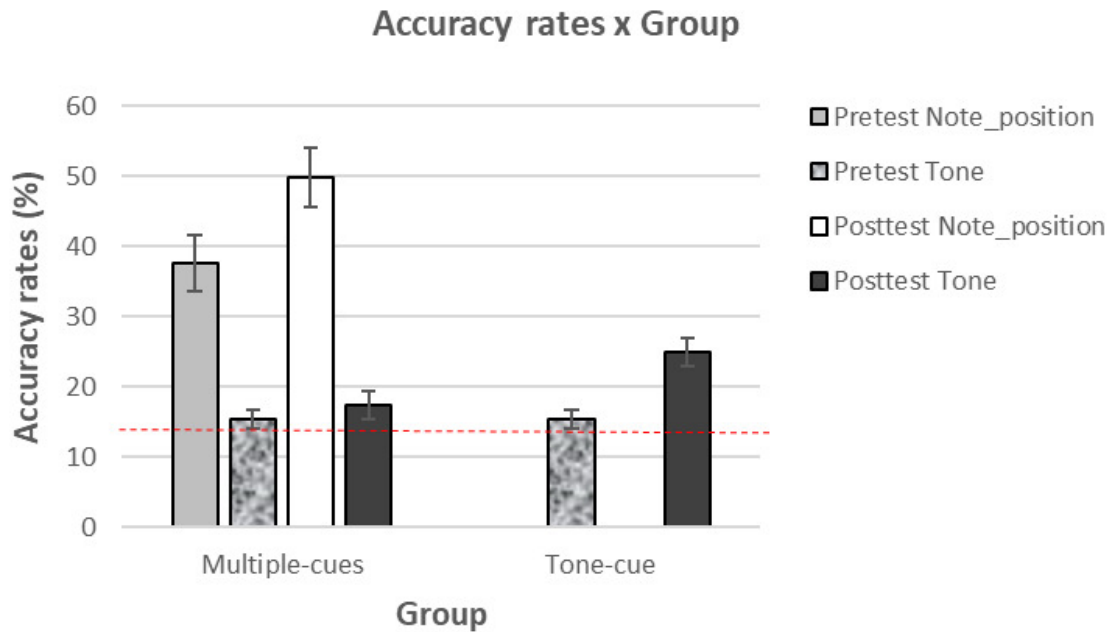


Figure 6. Difference in Accuracy rates between Group in Experiment 1

Note. Differences in Accuracy rates (pretest and posttest) between the groups (multiple-cues vs. tone-cue). Standard error bars and accuracy chance guessing at 14.3% (in red) are shown in the figure.

Discussion

In Experiment 1, we wanted to study whether nonmusicians were able to easily and rapidly learn pitch-label associations. Our results showed that, as expected, both groups of participants showed a contingency effect in the learning phase. We note that while contingency effects were larger in the multiple-cues group in both response times and errors, this finding should be interpreted with caution, as the multiple-cues group could be biased not only by the tones, but also by the predictive note positions. Also interesting, both groups were able to respond above chance in the tests phases in line with previous findings in the contingency learning literature (Iorio et al., 2023; Schmidt & De Houwer, 2019) and overall they significantly decreased their RTs in the posttest compared to the pretest. However, participants in the multiple-cues group seemed to show worse performance compared to the tone-cue group. Therefore, although previous research seems to suggest that presenting both note position and tones can benefit the learning of sub-skills (Mishra, 2014), our results suggest that when it comes to pitch identification presenting more than one predictive cue may interfere with the acquisition between the note name and the tone (i.e., an overshadowing effect).

Experiment 2

In Experiment 1, we studied the more experimentally “pure” case of nonmusicians learning to identify note pitches who, incidentally, have also clearly “missed” any potential critical period for acquiring pitch identification skills (see Introduction). This sort of sample would also cor-

respond to novice musicians just beginning to learn music. Learning to improve pitch identification skills could also be useful for experienced musicians. In that vein, Experiment 2 studies whether our incidental learning procedure can help musicians to improve their ability to identify and label tones. Incidentally, musicians are also an interesting group to study for another reason. Our participants were musicians but AP non-possessors. If pitch identification skills are strictly dependent on “good genes” (see Introduction), then this group seems to be the most unlikely to have said genes: they have had more than enough experience seeing music notation, playing notes, and hearing the corresponding pitches to have acquired AP already if they had the right disposition for it. Since previous research has suggested that AP development is related to early musical training (Crozier, 1997; Deutsch et al., 2006; Miyazaki & Ogawa, 2006), we also decided to take this measure into account as a covariate in our analysis.

As a further manipulation in Experiment 2, we introduced a second name-to-key mapping to be able to test for spatial compatibility effects. Specifically, Rusconi et al. (2006) suggested that the human cognitive system automatically codes pitches spatially with the highest pitches represented on the right and the lowest pitches on the left (akin the Spatial Musical Association of Response Codes, SMARC effect) and recent research indicates that space-pitch associations exhibit greater stability when supplemented with metaphors embedded in language (Dolscheid et al., 2020). The French language, for instance, expresses pitch predominantly in terms of spatial height. In Experiment 1, the lowest note name for the lowest pitch in our task, “fa”, corresponded to the leftmost key on the key-

board, “Z”. Possibly, this could help with the acquisition of key-label responses based on the research on the SMARC effect. As a small note, we did not find evidence for a SMARC effect on the acquisition of the note name/note position associations in our previous work (Iorio et al., 2023). However, to control for possible influence of the SMARC effect in our paradigm, we compared performance in two groups (compatible vs. incompatible groups, see the *Method* section for more details).

Our primary hypothesis for the present experiment was that musicians would be able to improve their pitch identification, similar to the nonmusicians in Experiment 1. Given that the tone-cue manipulation improved posttest note detection notably more than the multiple-cues manipulation in Experiment 1, we dropped the multiple-cues condition from Experiment 2. Additionally, we hypothesized that pitch identification abilities would be higher for participants that started learning music earlier on in life. To what extent early music learning might interact with pre/post improvement scores was uncertain.

Method

Participants

The recruitment process was similar to the one used in Experiment 1, except that we searched for musicians rather than nonmusicians. Therefore, as specified in the recruitment advertisement, we looked for French speaking participants with experience in playing music. 117 participants took part in the experiment and received monetary compensation (3.80 £) for their participation. However, 9 participants were excluded from the analysis because they failed to report information about the age they started musical training, information that we used as a covariate in the following analyses. Of the remaining 108 participants, 19 declared to have absolute pitch. However, only seven participants reported accuracy rates between 60% and 100% in the pretest and were discarded from the following analysis, as in Experiment 1. All participants accepted a written consent before beginning the study. All the procedures were conducted in accordance with the Declaration of Helsinki. Participants’ anonymization was guaranteed.

Apparatus, Design, and Procedure

The general structure of the experiment was similar to the one used in Experiment 1 with some exceptions. Firstly, we changed the name-to-key assignment (the keys D, F, G, H, J, K, L, instead of the keys Z,E,R,T,Y,U,I) to control for possible differences in the keyboards used by the par-

ticipants recruited online. As mentioned above, we introduced a second name-to-key mapping to be able to test for spatial compatibility effects. For the first group (compatible group), we used the same name-to-key assignment as for Experiment 1 (i.e., the tones used went from “fa” to “mi”, corresponding to the D to L keys on the keyboard). In this group, the spatial position of the tones was “compatible”, or in other words matched, the responses. For the second group (incompatible group), we used the D to L keys to refer to “do” to “si” note names. In this group, there was no spatial compatibility between the tones and the responses (e.g., the leftmost “D” key corresponded to one of the highest tones, viz., “do”). As one further change, we excluded the multiple-cues condition. All participants completed the tone-cue condition from Experiment 1.

Data Analysis

As in Experiment 1, we ran an ANOVA on RTs and error rates for the learning phase and t-tests on accuracy and RTs for the test phases. However, here we additionally added information about the start of musical training as a covariate in our analysis. 12.77% of the trials on the total number of participants were eliminated based on the same criteria used in Experiment 1 (i.e., trials in which participants failed to respond in 3000 ms). All analyses were evaluated at the $\alpha = .05$ level of significance, and we reported the Bayes factor. The data set is available via the following link: <https://osf.io/xjdt4/>.

Results

Response Times

The repeated measures ANOVA for RTs with the factors Contingency (high vs. low) and Group (compatible vs. incompatible) and the age of the start of musical training as covariate showed a significant main effect of Contingency, $F(1,98) = 36.97, p < .001, \eta^2 = .274, BF_{10} > 100$, indicating faster responses for high-contingency trials ($M = 860\text{ms}, SD = 209$) than for low-contingency trials ($M = 909\text{ms}, SD = 221$). The main effect of Group was not significant, $F(1,98) = 1.59, p = .210, \eta^2 = .016, BF_{10} = .702$. The interaction between Contingency and Group was significant (Figure 7), $F(1,98) = 4.12, p = .045, \eta^2 = .040, BF_{10} = 1.02$, due to a greater difference between high and low contingency trials for the compatible group (Mhigh-contingency= 878ms, $SD = 214$, Mlow-contingency trials = 942ms, $SD = 227$) compared to the incompatible group: (Mhigh-contingency= 840ms, $SD = 204$, Mlow-contingency trials = 872ms, $SD = 210$).

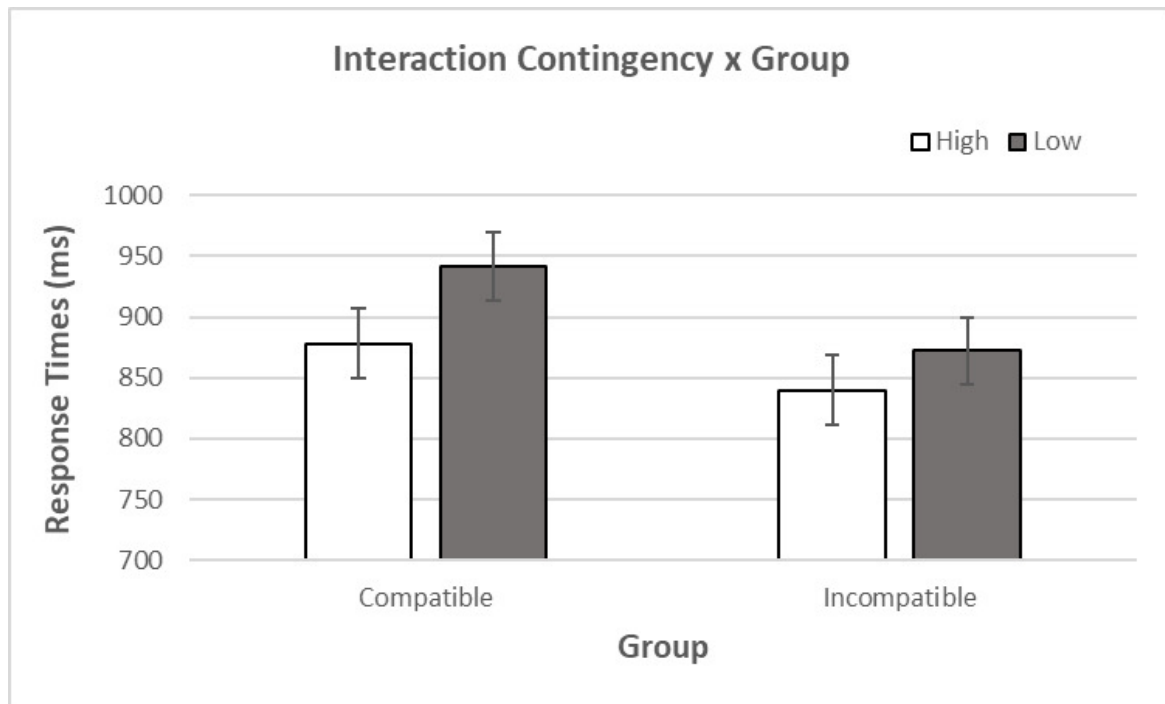


Figure 7. Interaction between contingency effect and Group in Experiment 2

Note. Interaction between Contingency (High and Low) and Group (compatible vs. incompatible), standard error bars are shown in the figure.

The contingency effect was significant for both the compatible group, $M_{\text{low-high}} = 63.7$, $SD = 86.6$; $t(52) = 5.36$, $p < .001$, $d = .736$, $BF_{10} > 100$, and the incompatible group, $M_{\text{low-high}} = 32.5$, $SD = 70.2$; $t(47) = 3.21$, $p = .002$, $d = .436$, $BF_{10} = 13.1$.² The interaction between Contingency and beginning of the musical training was not significant, $F(1,98) = .621$, $p = .433$, $\eta^2 = .006$, $BF_{10} = .301$.

Error Rates

The repeated measures ANOVA for errors with the factors Contingency (high vs. low) and Group (compatible vs. incompatible) revealed a main effect of Contingency, $F(1,99) = 17.51$, $p < .001$, $\eta^2 = .150$, $BF_{10} > 100$ (more errors for low trials $M = 15.3\%$, $SD = 11.8$, compared to high trials $M = 12.5\%$, $SD = 9.75$), and a non-significant main effect for Group, $F(1,99) = 1.27$, $p = .263$, $\eta^2 = .013$, $BF_{10} = .573$. The interaction between Contingency and Group was also not significant, $F(1,99) = 1.27$, $p = .262$, $\eta^2 = .013$, $BF_{10} = .382$.

Test Phases

71.29% of participants (77 of 108) noticed the contingencies between the tones and the note names. We performed an ANOVA on the accuracy rates with Test (pre vs. post) and Group (compatible vs. incompatible) as factors. The main effect of Test was significant, $F(1,99) = 27.80$, $p < .001$, $\eta^2 = .219$, $BF_{10} > 100$, as well as the main effect of Group, $F(1,99) = 17.4$, $p < .001$, $\eta^2 = .150$, $BF_{10} > 100$. Also the interaction Test x Group was significant, $F(1,99) = 7.70$, $p = .007$, $\eta^2 = .072$, $BF_{10} = 5.98$ (pre-test: $M_{\text{compatible}} = 17.6\%$, $SD = 14.5$; $M_{\text{incompatible}} = 11.4\%$, $SD = 10.1$. post-test: $M_{\text{compatible}} = 32.0\%$, $SD = 23.1$; $M_{\text{incompatible}} = 15.9\%$, $SD = 13.0$).

T-tests for pretest and posttest accuracy rates, shown in Figure 8, revealed that the compatible group did not perform significantly above chance (i.e., 14.3%) in the pretest, $t(52) = 1.66$, $p = .104$, $d = .228$, $BF_{10} = .537$, $M = 17.6\%$, $SD = 14.5$, but were significantly above chance in the posttest, $t(52) = 5.58$, $p < .001$, $d = .766$, $BF_{10} > 100$, $M = 32.0\%$, $SD = 23.1$. The improvement between pre/posttest was signif-

² An anonymous reviewer suggested that the compatibility effect might be particularly present for pianists, where the left-to-right assignment of pitches to keys is particularly salient. To explore this possibility, we separated participants into two groups: those who reported studying piano ($n = 37$) and those who reported studying an instrument other than piano or singing ($n = 64$). In an ANOVA including this extra factor, the main effect of Contingency remained significant, $F(1,97) = 35.55$, $p < .001$, $\eta^2 = .268$. The main effect of Group, $F(1,97) = 10.9$, $p = .298$, $\eta^2 = .011$, and Instrument played, $F(1,97) = 1.17$, $p = .282$, $\eta^2 = .012$, were not significant. Most critically, the interaction between Contingency x Group x Instrument played was not significant, $F(1,97) = 1.31$, $p = .254$, $\eta^2 = .013$, though we do note that there was at least a numerical trend for a larger Contingency by Group interaction for piano players (55 ms interaction; compatible group: $M_{\text{high-contingency}} = 826\text{ms}$, $SD = 168$, $M_{\text{low-contingency}} = 904\text{ms}$, $SD = 182$; incompatible group: $M_{\text{high-contingency}} = 829\text{ms}$, $SD = 191$, $M_{\text{low-contingency}} = 852\text{ms}$, $SD = 197$) relative to other instrument players (17 ms interaction; compatible group: $M_{\text{high-contingency}} = 905\text{ms}$, $SD = 232$, $M_{\text{low-contingency}} = 961\text{ms}$, $SD = 247$; incompatible group: $M_{\text{high-contingency}} = 847\text{ms}$, $SD = 215$, $M_{\text{low-contingency}} = 886\text{ms}$, $SD = 220$).

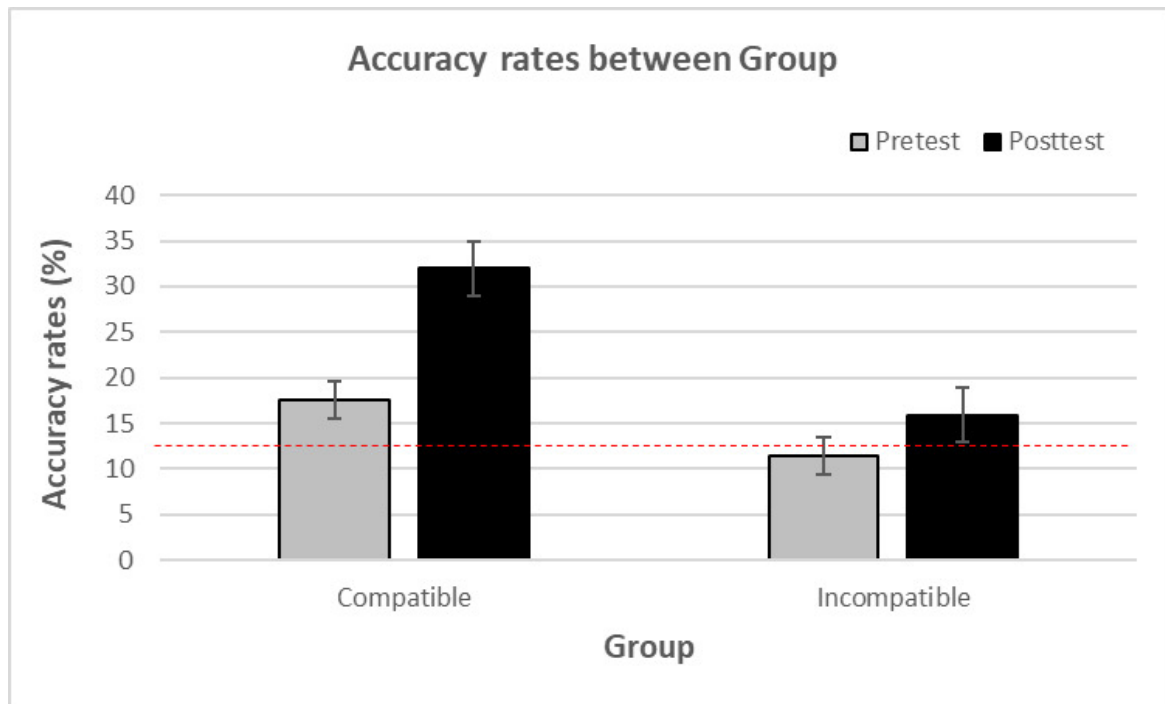


Figure 8. Difference in Accuracy rates between Group in Experiment 2

Note. Differences in Accuracy rates (Pretest and Posttest) between the groups (compatible vs. incompatible), standard error bars and accuracy chance guessing at 14.3% are shown in the figure.

icant for this group $t(52) = 5.01, p > .001, d = .688, BF_{10} > 100$.

The incompatible group did not perform above chance in the pretest, $t(47) = -1.974, p = .054, d = -.285, BF_{10} = .930, M = 11.4\%, SD = 10.1$, and in the posttest, $t(47) = .835, p = .408, d = .121, BF_{10} = .218, M = 15.9\%, SD = 13.0$. Although the incompatible group reported performance slightly below chance guessing in the pretest, their performance significantly improved between pre/posttest, $t(48) = 2.22, p = .031, d = .320, BF_{10} = 1.45$.³

We ran an ANOVA with the factors Test (pre vs. post) and Group (compatible vs. incompatible) on the RTs for the tone naming task. This analysis showed a nonsignificant main effect for Test, $F(1,99) = 3.834, p = .053, \eta^2 = .037, BF_{10} = 1.147$, a nonsignificant main effect of Group, $F(1,99) = .002, p = .959, \eta^2 = .000, BF_{10} = .175$, and a non-significant interaction between Test and Group, $F(1,99) = .353, p = .554, \eta^2 = .003, BF_{10} = .258$.

Discussion

In Experiment 2, we wanted to determine whether our incidental learning procedure could help musicians to improve their ability to identify and label tones. The results showed a significant contingency effect for both groups in response times and errors. However, our findings on

the RTs in the test phases did not reveal any general significant improvement in response times between the pre- and posttest. Furthermore, only the compatible group performed significantly above chance in the posttest, though both groups showed an increase in performance between the pre- and posttest. Similarly, the RT contingency effect was larger in the compatible group. These outcomes suggest that when asked to explicitly name a tone, participants may rely on some sort of internal spatially related code for tones, as shown in previous research (Ariga & Saito, 2019; Rusconi et al., 2006). Participants can learn the contingencies in either case, but spatial compatibility may help.

Experiment 3

In Experiment 3, we extend the results of the preceding experiments in two ways. First, we tested whether pitch learning persists over time. Considering how previous research about pitch identification describes the acquisition of pitch-label associations as something difficult, it might be the case that the learning effect we have shown in the Experiment 1 is just temporary and will not persist over time. That is, what if nonmusicians were able to learn the pitch-label associations as a result of the many repetitions they were exposed to, but they did not form any long-lasting representations of these associations in memory? Here

³ Again, we ran a non-parametric repeated measures ANOVA (Friedman test) on the tone naming accuracy. The improvement from pretest to posttest was significant in the compatible group, $\chi^2 = 16.5, p < .001$. For the incompatible group, the improvement was not significant, $\chi^2 = .641, p = .425$.

we argue that pitch-label associations can not only be incidentally learned, but that the information is retained in memory and it can be easily retrieved not only immediately after learning it, but even more interestingly after some time from the learning process. We therefore hypothesized that posttest scores would still be increased after a delay (i.e., that the learned pitch information remains in memory).

Second, we aimed to study the effect of intentionality on the learning and consolidation of pitch-label associations. Past research suggests that being aware of the contingencies before beginning the experiment benefits their acquisition (Schmidt & De Houwer, 2012b, 2012a, 2012c). That is, while participants who are not informed about the regularities in the task generally still learn said regularities, participants informed in advance about the contingency manipulation often show even larger learning effects. Similar results have also been observed in sequence learning studies (Destrebecqz, 2004). This instruction effect is often only moderate and is not always robust. For instance, in the above-mentioned sight-reading studies (Iorio et al., 2023) we did not find any significant differences in the learning phase between participants that were aware of the contingencies and those that incidentally learned them. Numerical differences were suggestive, however, and posttest ratings were improved with explicit instruction. To assess this question in the pitch learning context, we therefore created two groups: an incidental learning group that was not informed about the manipulation before starting the experiment and a deliberate learning group that was. We expected larger learning effects in the deliberate learning group relative to the incidental learning group, both in the performance measures during the learning phase and in the posttest scores. That is, it is possible that being attentive to the contingencies helps with consolidation more than learning in a purely incidental way.

Method

Participants

268 students from the University of Burgundy took part in this experiment. The experiment was part of a second-year cognitive psychology tutorial and served as the basis for student presentations. Students were not informed about the purpose of the experiment until after completing both phases, however. Due to complications with the COVID pandemic, the study was also conducted online using the same software as the preceding experiments (PsyToolkit; Stoet, 2010, 2016). We excluded participants that either did not complete all the test phases or did not correctly indicate their student number (which did not allow us to match their datasets together). 136 participants that

met these conditions and declared to not have AP were randomly divided into an incidental learning group (73 in total) and a deliberate group (63 in total). One participant was removed from the sample because accuracy was between 60% and 100% in the pretest. As in the previous studies, all participants signed a consent form before starting the study. The study was consistent with the Declaration of Helsinki and participants' anonymization was guaranteed.

Apparatus, Design, and Procedure

The experiment followed the same structure as the previous studies with some exceptions. First, the participants were divided into incidental and deliberate learning groups. While in the first group participants were not instructed about the contingencies (i.e., as in the prior experiments), in the deliberate group participants were told about the contingencies before beginning the experiment and they were encouraged to learn them, translated from French:

“Note: Each note will be presented more frequently with the correct tone and less frequently with the incorrect tones. Try to learn the note name for each tone.”

As an additional change, in order to study the consolidation of new material, we also added a (surprise) follow up session one week after the end of the learning phase. During the follow up, participants were asked to take part in a second posttest tone naming task, which was identical in all respects to the other posttest (and pretest). As a minor aside, we note that students were also asked to fill in a paper-and-pencil survey with various questions about their prior music experiences. We note that this survey was included for purely pedagogical purposes, and we have not nor had ever intended to analyze these data, with some exceptions for the questions used for controlling for musical expertise mentioned below.⁴

Data Analysis

The analysis was based on the same criteria as those used in Experiments 1 and 2. We conducted a repeated measures ANOVA for RTs with musical expertise as a covariate and for error rates to assess the overall main effects of Contingency, Group, and the interaction between them. Following the exclusion criteria used in Experiments 1 and 2, we discarded 13.37% of the data. We ran *t*-tests and ANOVAs on Accuracy and RTs for the test phases and the follow up. All analyses were evaluated at the $\alpha = .05$ level of significance. Again, the Bayes factor was reported for each analysis. The data set is available via the following link: <https://osf.io/xjdt4/>.

⁴ In fact, the surveys were printed the prior year for an unrelated study and had not been used due to the COVID pandemic (an electronic version was used instead). We decided to use these questionnaires both (a) because they contained a few questions related to our selection criteria, and (b) to give students inspiration for potential discussion points in their group presentations. The non-pertinent questions, however, have not been coded electronically.

Results

Response Times

We ran a repeated measures ANOVA for RTs with the factors Contingency (high vs. low) and Group (incidental vs. deliberate) that indicated a significant main effect of Contingency, $F(1,133) = 13.75$, $p < .001$, $\eta^2 = .094$, $BF_{10} = 45.26$, showing faster responses for high-contingency trials ($M = 933\text{ms}$, $SD = 164$) than for low-contingency trials ($M = 963\text{ms}$, $SD = 171$). The main effect of Group was not significant, $F(1,133) = .661$, $p = .418$, $\eta^2 = .005$, $BF_{10} = .459$. The interaction between Contingency and Group was not significant, $F(1,133) = 2.20$, $p = .140$, $\eta^2 = .016$, $BF_{10} = .489$, though there was a numerical trend towards a larger contingency effect for the deliberate learning group, $M_{low-high\ contingency\ trials} = 43.2$, $SD = 111$, compared to the incidental group, $M_{low-high\ contingency\ trials} = 18.5$, $SD = 81.3$.

Error Rates

A repeated measures ANOVA for errors with the factors Contingency (high vs. low) and Group (incidental and deliberate) revealed a significant main effect of Contingency, $F(1,133) = 16.74$, $p < .001$, $\eta^2 = .112$, $BF_{10} > 100$, showing more errors for low-contingency trials ($M = 15.4\%$, $SD = 10.9$) than for high-contingency trials ($M = 13.1\%$, $SD = 8.48$). The main effect of Group, $F(1,133) = .291$, $p = .590$, $\eta^2 = .002$, $BF_{10} = .367$, and the interaction between Contingency and Group, $F(1,133) = .040$, $p = .840$, $\eta^2 = .000$, $BF_{10} = .185$, were not significant (incidental group, $M_{high-trials} = 12.8\%$, $SD = 7.47$, $M_{low-trials} = 15.0\%$, $SD = 9.53$; deliberate group, $M_{high-trials} = 13.5\%$, $SD = 9.59$, $M_{low-trials} = 16.0\%$, $SD = 12.3$).

Test Phases

For the subjective awareness question, 59.25% (80 of 135) of the participants noticed the contingencies between tones and note names. The incidental learning group performed significantly above chance in the pretest, $t(72) = 2.29$, $p = .025$, $d = .268$, $BF_{10} = 1.49$, $M = 17.4\%$, $SD = 11.4$, in the posttest, $t(72) = 8.22$, $p < .001$, $d = .962$, $BF_{10} > 100$, $M = 29.9\%$,

$SD = 16.2$, and in the follow up, $t(72) = 4.54$, $p < .001$, $d = .531$, $BF_{10} > 100$, $M = 22.6\%$, $SD = 15.7$. For the deliberate group, performance was not significantly above chance in the pretest, $t(61) = 1.50$, $p = .139$, $d = .191$, $BF_{10} = .402$, $M = 17.0\%$, $SD = 14.0$, but was significant in the posttest, $t(61) = 6.75$, $p < .001$, $d = .857$, $BF_{10} > 100$, $M = 30.3\%$, $SD = 18.7$, and in the follow up, $t(61) = 5.68$, $p < .001$, $d = .721$, $BF_{10} > 100$, $M = 28.7\%$, $SD = 20.0$. The differences in accuracy rates found between the groups were not significant in the

pretest, $t(133) = .173$, $p = .863$, $d = .029$, $BF_{10} = .187$, or in the posttest, $t(133) = -.154$, $p = .878$, $d = -.026$, $BF_{10} = .187$, however the deliberate group performed significantly better than the incidental group in the follow up, $t(133) = 1.981$, $p = .0496$, $d = .342$, $BF_{10} = 1.091$.

We performed an ANOVA⁵ on the accuracy rates with Test (pre vs. post) and Group (deliberate vs. incidental) as factors. The main effect of Test was significant, $F(2,266) = 41.49$, $p < .001$, $\eta^2 = .238$, $BF_{10} > 100$. The main effect of Group was not significant, $F(1,133) = .852$, $p = .358$, $\eta^2 = .006$, $BF_{10} = .260$. The interaction Test x Group was also not significant, $F(2,266) = 2.96$, $p = .053$, $\eta^2 = .022$, $BF_{10} = .710$ (pre-test: $M_{incidental} = 17.4\%$, $SD = 11.4$; $M_{deliberate} = 17.0\%$, $SD = 14.0$; post-test: $M_{incidental} = 29.9\%$, $SD = 16.2$; $M_{deliberate} = 30.3\%$, $SD = 18.7$; follow up: $M_{incidental} = 22.6\%$, $SD = 15.7$; $M_{deliberate} = 28.7\%$, $SD = 20.0$).

Accuracy rates were significantly higher in posttest compared to pretest in both groups, as shown in Figure 9: incidental group, $t(72) = 5.99$, $p < .001$, $d = .701$, $BF_{10} > 100$, deliberate group, $t(61) = 5.95$, $p < .001$, $d = .711$, $BF_{10} > 100$. Accuracy rates were significantly lower in the follow-up compared to the posttest in the incidental group, $t(72) = -4.34$, $p < .001$, $d = -.508$, $BF_{10} > 100$, and not significantly different for the deliberate group, $t(61) = .806$, $p = .423$, $d = .102$, $BF_{10} = .190$. Most importantly, accuracy rates were significantly higher in the follow up compared to the pretest for both groups: incidental group, $t(72) = 3.12$, $p = .003$, $d = .365$, $BF_{10} = 10.6$, and deliberate group, $t(61) = 4.84$, $p < .001$, $d = .615$, $BF_{10} > 100$.

In the ANOVA with the factors of Test (pre and post) and Group (incidental and deliberate) on response times, while the main effect of Test was significant, $F(2,266) = 10.86$, $p < .001$, $\eta^2 = .076$, $BF_{10} > 100$, showing a decrease in response times between the tests ($M_{pretest} = 2686$ ms, $SD = 3081$; $M_{posttest} = 1816$ ms, $SD = 1223$; $M_{follow\ up} = 1677$ ms, $SD = 1016$), both the main effect for Group, $F(1,133) = .018$, $p = .892$, $\eta^2 = .000$, $BF_{10} = .131$, and the interaction between Test and Group, $F(2,266) = .175$, $p = .840$, $\eta^2 = .001$, $BF_{10} = .060$, were nonsignificant.

Discussion

The results of Experiment 3 showed an overall significant main effect of Contingency in both groups. Despite the larger sample size and contrary to hypotheses, a nonsignificant interaction between Contingency and Group indicated no clear evidence that intentionality helps the acquisition of the contingency during this performance task. Similarly, no effect of intentionality was observed in the posttest scores, both immediately and one week after the learning phase. There were some hints of larger learning effects, at least in the response times, but overall deliberate

5 Non-parametric repeated measures ANOVAs (Friedman test) on the accuracy rates revealed significant improvement from pretest to posttest in the incidental, $\chi^2 = 16.0$, $p < .001$, and deliberate group, $\chi^2 = 19.8$, $p < .001$. The difference between pretest and follow up was also significant for the incidental, $\chi^2 = 8.14$, $p = .004$, and deliberate group, $\chi^2 = 11.0$, $p < .001$. Finally, the decrease between posttest and follow up was significant for the incidental group, $\chi^2 = 7.56$, $p < .006$, but not for the deliberate group, $\chi^2 = .276$, $p = .599$.

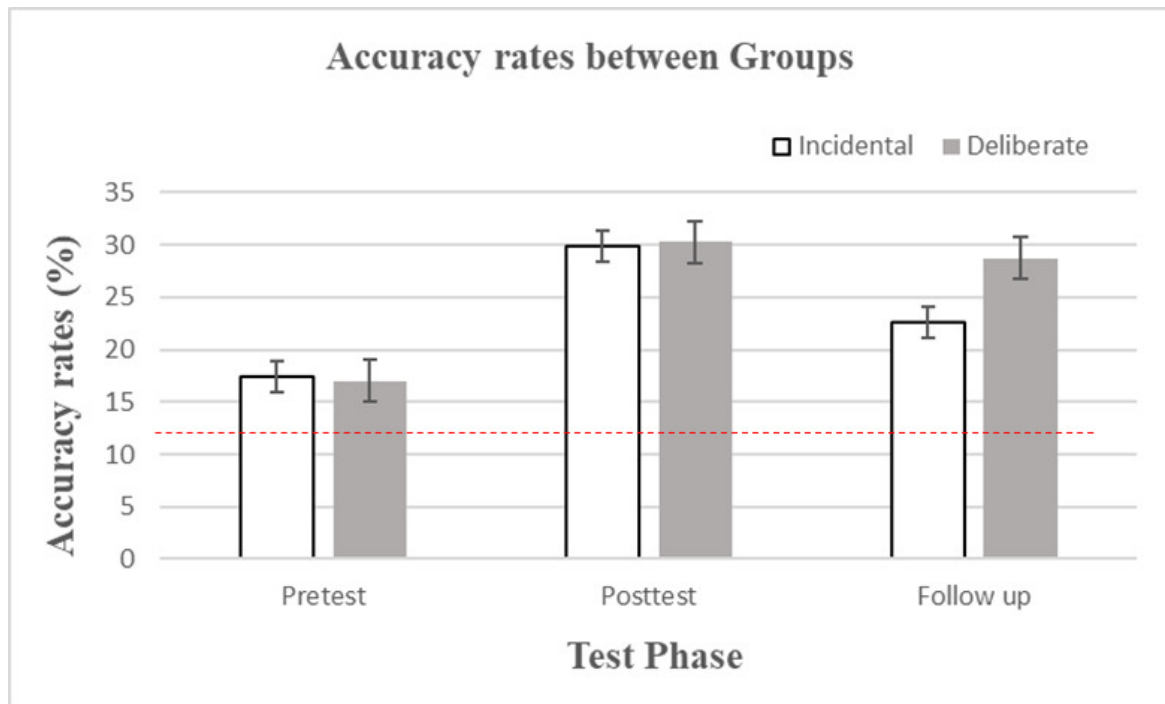


Figure 9. Difference in Accuracy rates between Group in Experiment 3

Note. Differences in Accuracy rates (Pretest, Posttest and Follow-up) between the groups (Incidental vs. Deliberate), standard error bars and accuracy chance guessing at 14.3% are shown in the figure.

learning did not seem to increase learning effects drastically. Instead, learning effects were robust in all phases of the experiment for both groups, including the one week follow-up posttest. These results may reflect an important role played by incidental learning for the internalization of pitch-label associations.

General Discussion

The current research investigated the learnability of pitch names for auditory pitch stimuli in an incidental learning task. As mentioned in the Introduction, the goal of the present work was not to investigate the learnability of absolute pitch, strictly defined, but rather to explore whether there is a fundamental difficulty in learning pitch-label associations. Our results support the idea that pitch-label associations are learnable in incidental learning conditions. Indeed, learning effects were observed during a very short learning procedure. In Experiment 1, the results indicated that nonmusicians were able to incidentally learn pitch-label associations and properly use this information to name tones above chance guessing in a tone naming task. The multiple-cues group, which was exposed to the combination of note positions and tones as the cues during the learning phase, were less accurate in the posttest compared to the tone-cue group. As previously mentioned, this result may be due to the well-known overshadowing effect. That is, if two stimuli, A and X (or in this specific case, the note position and the tone), are presented together and are followed by an outcome (the note name in our study), learning about the relation between X and the outcome is often weaker compared to when only stimulus X is paired

with the outcome (Kamin, 1969; Pavlov, 1927). The data are thus consistent with overshadowing, given that the multiple-cues group performed more poorly than the tone-cue group. Furthermore, these results seem to be inconsistent with the idea that combining auditory and visual information boosts musical learning, as previously suggested in a sight-reading context (see Mishra, 2014 for a review). Of course, there are both auditory and visual information that are important to learn in music learning, but combining the two into one learning procedure may be suboptimal. We briefly note that there are several competing theories of overshadowing and other cue competition phenomena. For instance, the Rescorla-Wagner (1972) model suggests that learning of one association (e.g., between pitches and note names in the current experiment) is impaired to the extent that another association (e.g., between note positions and note names) is strong enough to anticipate the outcome (e.g., note name), at which point associations are less strongly updated. According to another view (Mackintosh, 1975), attention is drawn to a more salient stimulus (e.g., note positions) which reduces learning for the “overshadowed” stimulus (e.g., pitches). Regardless of the exact mechanism, our results suggest that presenting musical notation does not help with learning to identify pitches by ear.

In Experiment 2, we further investigated the efficacy of an incidental learning procedure in improving pitch identification in participants with previous musical experience. Similar to the results for nonmusicians, musicians were also able to strengthen their knowledge about pitch-label associations and use this information to correctly guess above chance the name of the tones in the posttest tone

naming task. In addition, in Experiment 2 we controlled for the possible influence of the SMARC effect (Ariga & Saito, 2019; Rusconi et al., 2006) on pitch-label acquisition. In the compatible group, tones were spatially congruent with the position of the keys on the keyboard, whereas in the incompatible group the leftmost tone “mi” was mapped to one of the rightmost keys on the keyboard. Surprisingly, the incompatible group responded faster than the compatible group in the learning phase. The reason for this is unclear but might be related to the fact that the response keys for the incompatible group were ordered from “do” to “si”, a more “classical” order of the note names (i.e., the order that most learn in elementary school). This may have facilitated overall RTs. However, spatial compatibility did influence test phase performance in the anticipated direction. Accuracy in the compatible group improved significantly from pre- to post-test, but this improvement was much smaller in the incompatible group. This latter result is coherent with the notion that we spatially code pitches (Ariga & Saito, 2019; Rusconi et al., 2006): the incongruity between pitches and the spatial location in the incompatible group may have interfered with the more natural codes and therefore negatively influenced the acquisition of the pitch-label associations. Pitch learning clearly occurred (i.e., given the pre-post improvements), but spatial incompatibility seems to make this learning more difficult.

In Experiment 2, we also measured the age at which participants began musical training, which previous research suggests may have an impact on the internalization of pitches (Crozier, 1997; Deutsch et al., 2006; Miyazaki & Ogawa, 2006). Surprisingly, our results did not reveal any influence of this factor on the contingency effect (i.e., age of beginning music training did not interact with contingency). On the contrary, these results seem to point to the idea that even those who started musical training later than the critical period (i.e., between 4 and 5 years old) can still improve their performance in the auditory domain, suggesting the presence of a changeable internal pitch representation rather than a stable “pitch template”.

Finally, in Experiment 3 we focused on studying the role of incidental learning in the acquisition and consolidation of pitch-label associations in longer-term memory. Once again, as already reported in Experiment 1, nonmusicians showed significant contingency effects in both an incidental and a deliberate learning group. However, no notable differences were observed in the size of these learning effects, suggesting that being aware of the contingencies does not necessarily help to learn them better in performance tasks (or at least not to a substantial degree). On the other hand, we did find some differences in performance between the two groups in the test phases. The deliberate group not only reported higher accuracy rates (although the difference in accuracy rates between the groups was not significant) in the posttest compared to the incidental group, they also performed better in the follow up. In line with previous research (Iorio et al., 2023; Schmidt & De Houwer, 2012c), these results may indicate that being attentive to the contingencies benefits the consolidation of the information acquired. However, when it comes to skill

automatization (e.g., as measured by RTs and error rates), it seems that intentionality does not positively increase performance substantially.

Limitations and Future Directions

It is important to reiterate that the goals of the present work diverge from those of past work on pitch identification learning and, more particularly, learning of AP. As mentioned in the Introduction, much work on this topic has focused on the determinants of absolute pitch, with both genetic factors and early music learning being indicated as key factors. Some debate has raged about whether absolute pitch (i.e., to a strict criterion) is learnable at all in the absence of early music training and/or the right genetic background. Although some studies have certainly indicated that improvements are possible with extended, focal training regimes, doubt persists as to whether it would be possible, for instance, for an adult with no prior music training to develop absolute pitch. Our work asked a notably different question: whether the same sort of rapid learning and automatization observed in (non-musical) incidental learning procedures can also be observed in a pitch learning context. That is, with a drastically shorter learning procedure, can evidence of improvements already be observed in explicit identification (i.e., in our test phases)? And similarly, do we see automatic biases on performance during learning? The answer to both of these questions seems to be “yes”. Posttest scores (accuracy and RTs), in addition to RTs and error rate indices during learning, are impacted by the acquired contingencies. In other words, participants not only improve their explicit tone naming scores, but this retrieval is fast and automatic. Overall, the results suggest that an incidental learning procedure can benefit the internalization of pitches and one reason why an incidental learning procedure like ours works may be because of the many repetitions that participants can experience in a small amount of time.

The present research may raise one interesting question. Why does our procedure work at all? Specifically, if it is possible to incidentally learn pitch-label associations rapidly, then why do all musicians not already have absolute pitch? Indeed, musicians spend many years playing and listening to notes and they know the names of said notes. One significant difference lies in the utilization of incidental learning, which typically is not emphasized in traditional musical training such as solfège exercises that usually involve deliberate, instructed training. Indeed, traditional musical instruction typically does not focus on specifically learning the associations between pitches and note names. For example, most musical practice involves learning procedural actions on the instrument from music notation, without the (necessary) intermediary of note names. Subsequent repetition involves repetition from procedural memory rather than specific practice of associating a pitch to a note name. Further, as suggested by an anonymous reviewer, this sort of music practice could also provoke overshadowing as the focus is *not* exclusively on the association between the note name and the pitch (e.g., also on the musical notation, which produced overshadowing in our Experiment 1). Even

aural exercises that are typical in traditional instruction (e.g., interval training) tend not to focus on *absolute* pitch detection. In general, traditional instruction (typically) does not seem to involve the type of learning that is highly analogous to the current task. The fact that learning is incidental is perhaps at least partially relevant, too. A key feature of more implicit types of learning is the rapidity of this learning. It also seems particularly effective in cases where a regularity is very difficult to learn in a conscious and deliberate manner. In some cases, intentional learning can actually *hurt* performance (Berry & Broadbent, 1988; Reber, 1976; Reber et al., 1980; Wulf et al., 1998). On the other hand, the results of our Experiment 3 do not suggest that deliberate learning hurts in the present task.

What remains to be explored, however, is whether this type of approach could be effective (e.g., with much longer training regimes, similar to past research) to acquire absolute pitch, and whether our approach is more effective than other alternatives. Indeed, the present work might pose an interesting question: Could our procedure (or something similar) be used to train an adult AP non-possessor to acquire AP, strictly defined, or something approaching this? The present experiments, while they do demonstrate impressively rapid improvements, do not speak directly to this question. Here, we consider why they do not and what future research might be conducted to empirically evaluate this question. First, we note that we presented participants with a relatively small number of notes, the seven notes of a C Major scale. We did not attempt to train participants with all 12 semitones of an octave (i.e., the smallest distance between notes in the Western scale). This is, of course, quite different than prior work investigating AP acquisition (Van Hedger et al., 2019; Wong, Lui, et al., 2020; Wong, Ngan, et al., 2020). The reason for this methodological choice was that, contrarily to the previous studies, we recruited nonmusicians and focused on short- and medium-term improvements in pitch identification. Additionally, to parallel the structure of a more typical incidental learning task, seven response choices is already quite a lot. Further, for a key press task, 12 responses would be further complicated by the fact that participants do not, obviously, have 12 fingers. Incidentally, in ongoing work with another graduate student of the last author using a conceptually similar procedure, we have observed similar improvements in explicit pitch naming with the 12 notes of an octave with a similarly short training procedure, though automatic performance during learning (e.g., response times) was not assessed in this ongoing research.

As another limitation, we did not implement standardized AP tests (see Van Hedger et al., 2019, for some examples of AP tests) in our work. Standardized tests, though varying in nature from one test to another, typically involve all 12 semitones of two or more octaves, often in multiple timbres (i.e., played by different instruments). These tests also frequently include large jumps of more than an octave between adjacent notes and potentially distracting white noise between notes to prevent RP-type strategies. It is therefore possible that participants in our experiments may not have learned pitch classes (e.g., the ability to identify

a C in any octave or timbre), but rather the pitch names of particular auditory stimuli. It is similarly possible that participants used some form of an RP-comparison strategy. We did not use this type of standardized test in the present research for a few reasons. First, as already mentioned, it was not our aim to claim that our procedure can teach AP to nonmusicians. Rather, it was the goal to determine whether there is something fundamentally unlearnable about pitch-label associations. Second, we aimed to study short- to medium-term learning in naive participants. Octave- and timbre-generalization in more stringent tests of AP, we imagine, would require longer training periods. In some recent and ongoing follow-up work, however, we have already observed transfer of learning from trained timbres to untrained timbres (Henry & Schmidt, 2023) and some initial positive results in octave generalization. However, future research with such standardized tests (and most likely: longer training) might aim to evaluate whether an approach like the current one is capable of producing true AP and how this might compare to other approaches.

In conclusion, in this series of studies we explored whether a musical contingency learning procedure could aid in the rapid acquisition and consolidation of pitch-label associations in memory. Although, our results suggest that incidental learning may have a positive role in the acquisition of pitch-label associations as well as on its consolidation, more research is needed in order to further determine the role of this kind of incidental acquisition in the auditory domain.

Authors Contributions

Contributed to conception and design: CI, JRS
 Contributed to acquisition of data: CI
 Contributed to analysis and interpretation of data: CI, JRS
 Drafted and/or revised the article: CI, EB, JRS
 Approved the submitted version for publication: CI, EB, JRS

Acknowledgements

This work was supported by the French “Investissements d’Avenir” program, project ISITEBFC (contract ANR15-IDEX-0003) to James R. Schmidt.

Competing Interests

We, the authors, declare that we have no conflicts of interest to disclose regarding the research, authorship, and/or publication of this manuscript. We affirm that we have no financial, personal, or other relationships with other people or organizations that could inappropriately influence, or be perceived to influence, our work.

Supplementary Material

Supplemental materials are available following the OSF link included in the manuscript. All the supplemental material will be openly available after publication.

Data Accessibility Statement

The data set is available via the following link: <https://osf.io/xjdt4/>.

Submitted: November 13, 2023 PDT, Accepted: May 10, 2024 PDT



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

References

- Akiva-Kabiri, L., & Henik, A. (2012). A unique asymmetrical Stroop effect in absolute pitch possessors. *Experimental Psychology*, *59*(5), 272–278. <https://doi.org/10.1027/1618-3169/a000153>
- Ariga, A., & Saito, S. (2019). Spatial–musical association of response codes without sound. *Quarterly Journal of Experimental Psychology*, *72*(9), 2288–2301. <https://doi.org/10.1177/1747021819838831>
- Armitage, J., & Eerola, T. (2020). Reaction time data in music cognition: Comparison of pilot data from lab, crowdsourced, and convenience web samples. *Frontiers in Psychology*, *10*. <https://doi.org/10.3389/fpsyg.2019.02883>
- Athos, E. A., Levinson, B., Kistler, A., Zemansky, J., Bostrom, A., Freimer, N., & Gitschier, J. (2007). Dichotomy and perceptual distortions in absolute pitch ability. *Proceedings of the National Academy of Sciences*, *104*(37), 14795–14800. <https://doi.org/10.1073/pnas.0703868104>
- Bermudez, P., & Zatorre, R. J. (2009). A distribution of absolute pitch ability as revealed by computerized testing. *Music Perception*, *27*(2), 89–101. <https://doi.org/10.1525/mp.2009.27.2.89>
- Berry, D. C., & Broadbent, D. E. (1988). Interactive tasks and the implicit-explicit distinction. *British Journal of Psychology*, *79*, 251–272. <https://doi.org/10.1111/j.2044-8295.1988.tb02286.x>
- Bianco, R., Harrison, P. M., Hu, M., Bolger, C., Picken, S., Pearce, M. T., & Chait, M. (2020). Long-term implicit memory for sequential auditory patterns in humans. *eLife*, *9*, e56073. <https://doi.org/10.7554/eLife.56073>
- Bigand, E., Perruchet, P., & Boyer, M. (1998). Implicit learning of an artificial grammar of musical timbres. *Current Psychology of Cognition*, *17*(3), 577–601.
- Bigand, E., & Poulin-Charronnat, B. (2006). Are we “experienced listeners”? A review of the musical capacities that do not depend on formal musical training. *Cognition*, *100*(1), 31. <https://doi.org/10.1016/j.cognition.2005.11.007>
- Bly, B. M., Carrión, R. E., & Rasch, B. (2009). Domain-specific learning of grammatical structure in musical and phonological sequences. *Memory & Cognition*, *37*(1), 10–20. <https://doi.org/10.3758/MC.37.1.10>
- Brandon, M., Terry, J., Stevens, C. (Kate), & Tillmann, B. (2012). Incidental learning of temporal structures conforming to a metrical framework. *Frontiers in Psychology*, *3*. <https://doi.org/10.3389/fpsyg.2012.00294>
- Carlson, K. A., & Flowers, J. H. (1996). Intentional versus unintentional use of contingencies between perceptual events. *Perception & Psychophysics*, *58*(3), 460–470. <https://doi.org/10.3758/BF03206821>
- Cheesman, J., & Merikle, P. M. (1984). Priming with and without awareness. *Perception & Psychophysics*, *36*(4), 387–395. <https://doi.org/10.3758/bf03202793>
- Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, *2*(10), 406–416. [https://doi.org/10.1016/S1364-6613\(98\)01232-7](https://doi.org/10.1016/S1364-6613(98)01232-7)
- Crozier, J. B. (1997). Absolute pitch: Practice makes perfect, the earlier the better. *Psychology of Music*, *25*(2), 110–119. <https://doi.org/10.1177/0305735697252002>
- De Houwer, J., & Beckers, T. (2010). A review of recent developments in research and theories on human contingency learning. *The Quarterly Journal of Experimental Psychology: Section B*, *55*(4), 289–310. <https://doi.org/10.1080/02724990244000034>
- Destrebecqz, A. (2004). The effect of explicit knowledge on sequence learning: A graded account. *Psychologica Belgica*, *44*(4), 217–247. <https://doi.org/10.5334/pb-44-4-217>
- Deutsch, D. (2013). Absolute Pitch. In *The Psychology of Music* (pp. 141–182). Elsevier. <https://doi.org/10.1016/B978-0-12-381460-9.00005-5>
- Deutsch, D., Henthorn, T., Marvin, E., & Xu, H. (2006). Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period. *The Journal of the Acoustical Society of America*, *119*(2), 719–722. <https://doi.org/10.1121/1.2151799>
- Dolscheid, S., Çelik, S., Erkan, H., Küntay, A., & Majid, A. (2020). Space-pitch associations differ in their susceptibility to language. *Cognition*, *196*, 104073. <https://doi.org/10.1016/j.cognition.2019.104073>
- Doorn, J. van, Bergh, D. van den, Böhm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharský, Š., Ly, A., Marsman, M., Matzke, D., Gupta, A. R. K. N., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E.-J. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review*, *28*(3), 813–826. <https://doi.org/10.3758/s13423-020-01798-5>
- Forrin, N. D., & MacLeod, C. M. (2017). Relative speed of processing determines color–word contingency learning. *Memory & Cognition*, *45*(7), 1206–1222. <https://doi.org/10.3758/s13421-017-0721-4>
- Grégoire, L., Perruchet, P., & Poulin-Charronnat, B. (2013). The Musical Stroop Effect. *Experimental Psychology*, *60*(4), 269–278. <https://doi.org/10.1027/1618-3169/a000197>
- Hamers, J. F., & Lambert, W. E. (1972). Bilingual interdependencies in auditory perception. *Journal of Verbal Learning and Verbal Behavior*, *11*(3), 303–310. [https://doi.org/10.1016/S0022-5371\(72\)80091-4](https://doi.org/10.1016/S0022-5371(72)80091-4)
- Henry, W., & Schmidt, J. R. (2023). *Stroop-like effects in pitch identification training and generalization to untrained timbres: Evidence from a contingency learning task* [Manuscript submitted for publication].

- Iorio, C., Šaban, I., Poulin-Charronnat, B., & Schmidt, J. R. (2023). Incidental learning in music reading: The music contingency learning task. *Quarterly Journal of Experimental Psychology*, 76, 429–449. <https://doi.org/10.1177/17470218221092779>
- JASP Team. (2019). *JASP (Version 0.11.1.0)*.
- Kamin, L. J. (1969). Predictability, surprise, attention and coding. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279–296).
- Kerka, S. (2000). *Incidental Learning. Trends and Issues Alert No. 18*. ERIC Publication. <https://eric.ed.gov/?id=ED446234>
- Leboe, L. C., & Mondor, T. A. (2007). Item-specific congruency effects in nonverbal auditory Stroop. *Psychological Research*, 71(5), 568–575. <https://doi.org/10.1007/s00426-006-0049-3>
- Levin, Y., & Tzelgov, J. (2016). Contingency learning is not affected by conflict experience: Evidence from a task conflict-free, item-specific Stroop paradigm. *Acta Psychologica*, 7. <https://doi.org/10.1016/j.actpsy.2015.12.009>
- Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics*, 56(4), 414–423. <https://doi.org/10.3758/BF03206733>
- Levitin, D. J. (2007). *This is your brain on music: The science of a human obsession*. Plume. <http://site.ebrary.com/id/10130584>
- Levitin, D. J., & Rogers, S. E. (2005). Absolute pitch: Perception, coding, and controversies. *Trends in Cognitive Sciences*, 9(1), 26–33. <https://doi.org/10.1016/j.tics.2004.11.007>
- Lewicki, P. (1985). Nonconscious biasing effects of single instances on subsequent judgments. *Journal of Personality and Social Psychology*, 48(3), 563–574. <https://doi.org/10.1037/0022-3514.48.3.563>
- Lewicki, P. (1986). Processing information about covariations that cannot be articulated. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12(1), 135–146. <https://doi.org/10.1037/0278-7393.12.1.135>
- Lewicki, P., Hill, T., & Czyzewska, M. (1992). Nonconscious acquisition of information. *American Psychologist*, 47(6), 796–801. <https://doi.org/10.1037/0003-066X.47.6.796>
- Loui, P., Wu, E. H., Wessel, D. L., & Knight, R. T. (2009). A generalized mechanism for perception of pitch patterns. *Journal of Neuroscience*, 29(2), 454–459. <https://doi.org/10.1523/JNEUROSCI.4503-08.2009>
- Mackintosh, N. J. (1975). A theory of attention: Variations in associability of stimuli with reinforcement. *Psychological Review*, 82(4), 276–298. <https://doi.org/10.1037/h0076778>
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109(2), 163–203. <https://doi.org/10.1037/0033-2909.109.2.163>
- MacLeod, C. M. (2019). Learning simple associations. *Canadian Psychology/Psychologie Canadienne*, 60(1), 3–13. <https://doi.org/10.1037/cap0000170>
- McKelvie, S. J. (1987). Learning and awareness in the Hebb digits task. *The Journal of General Psychology*, 114(1), 75–88. <https://doi.org/10.1080/00221309.1987.9711057>
- Miller, J. (1987). Priming is not necessary for selective-attention failures: Semantic effects of unattended, unprimed letters. *Perception & Psychophysics*, 41(5), 419–434. <https://doi.org/10.3758/BF03203035>
- Mishra, J. (2014). Improving sightreading accuracy: A meta-analysis. *Psychology of Music*, 42(2), 131–156. <https://doi.org/10.1177/0305735612463770>
- Miyazaki, K. (1988). Musical pitch identification by absolute pitch possessors. *Perception & Psychophysics*, 44(6), 501–512. <https://doi.org/10.3758/BF03207484>
- Miyazaki, K. (1990). The speed of musical pitch identification by absolute-pitch possessors. *Music Perception*, 8(2), 177–188. <https://doi.org/10.2307/40285495>
- Miyazaki, K., Makomaska, S., & Rakowski, A. (2012). Prevalence of absolute pitch: A comparison between Japanese and Polish music students. *The Journal of the Acoustical Society of America*, 132(5), 3484–3493. <https://doi.org/10.1121/1.4756956>
- Miyazaki, K., & Ogawa, Y. (2006). Learning absolute pitch by children: A cross-sectional study. *Music Perception*, 24(1), 63–78. <https://doi.org/10.1525/mp.2006.24.1.63>
- Miyazaki, K., & Rakowski, A. (2002). Recognition of notated melodies by possessors and nonpossessors of absolute pitch. *Perception & Psychophysics*, 64(8), 1337–1345. <https://doi.org/10.3758/BF03194776>
- Mordkoff, J. T., & Halterman, R. (2008). Feature integration without visual attention: Evidence from the correlated flankers task. *Psychonomic Bulletin & Review*, 15(2), 385–389. <https://doi.org/10.3758/PBR.15.2.385>
- Musen, G., & Squire, L. R. (1993). Implicit learning of color-word associations using a Stroop paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(4), 789–798. <https://doi.org/10.1037/0278-7393.19.4.789>
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19(1), 1–32. [https://doi.org/10.1016/0010-0285\(87\)90002-8](https://doi.org/10.1016/0010-0285(87)90002-8)
- Oberauer, K., Jones, T., & Lewandowsky, S. (2015). The Hebb repetition effect in simple and complex memory span. *Memory & Cognition*, 43(6), 852–865. <https://doi.org/10.3758/s13421-015-0512-8>
- Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex*. Oxford University Press.
- Pothos, E. M. (2007). Theories of artificial grammar learning. *Psychological Bulletin*, 133(2), 227–244. <https://doi.org/10.1037/0033-2909.133.2.227>
- Reber, A. S. (1967). Implicit Learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6(6), 855–863. [https://doi.org/10.1016/S0022-5371\(67\)80149-X](https://doi.org/10.1016/S0022-5371(67)80149-X)

- Reber, A. S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 88–94.
- Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, 118(3), 219–235. <https://doi.org/10.1037/0096-3445.118.3.219>
- Reber, A. S., Kassin, S. M., Lewis, S., & Cantor, G. (1980). On the relationship between implicit and explicit modes in the learning of a complex rule structure. *Journal of Experimental Psychology: Human Learning and Memory*. <https://doi.org/10.1037/0278-7393.6.5.492>
- Refaat, M. (2014). *A reaction time investigation of absolute pitch* [Ph.D., University of Warwick]. <http://wrap.warwick.ac.uk/60499/>
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.
- Rohrmeier, M., & Cross, I. (2009). Tacit tonality: Implicit learning of context-free harmonic structure. In *ESCOM 2009: 7th Triennial Conference of European Society for the Cognitive Sciences of Music* (p. 10).
- Rohrmeier, M., & Rebuschat, P. (2012). Implicit learning and acquisition of music. *Topics in Cognitive Science*, 4(4), 525–553. <https://doi.org/10.1111/j.1756-8765.2012.01223.x>
- Rusconi, E., Kwan, B., Giordano, B. L., Umiltà, C., & Butterworth, B. (2006). Spatial representation of pitch height: The SMARC effect. *Cognition*, 99(2), 113–129. <https://doi.org/10.1016/j.cognition.2005.01.004>
- Saffran, J. R., Aslin, R. N., & work(s);, E. L. N. R. (1996). Statistical Learning by 8-Month-Old Infants. *Science, New Series*, 274(5294), 1926–1928. <https://doi.org/10.1126/science.274.5294.1926>
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70(1), 27–52. [https://doi.org/10.1016/S0010-0277\(98\)00075-4](https://doi.org/10.1016/S0010-0277(98)00075-4)
- Saffran, J. R., Loman, M. M., & Robertson, R. R. W. (2000). Infant memory for musical experiences. *Cognition*, 77(1), B15–B23. [https://doi.org/10.1016/S0010-0277\(00\)00095-0](https://doi.org/10.1016/S0010-0277(00)00095-0)
- Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, 8(2), 101–105. <https://doi.org/10.1111/j.1467-9280.1997.tb00690.x>
- Salidis, J. (2001). Nonconscious temporal cognition: Learning rhythms implicitly. *Memory & Cognition*, 29(8), 1111–1119. <https://doi.org/10.3758/BF03206380>
- Schellenberg, E. G., & Trehub, S. E. (2003). Good Pitch Memory Is Widespread. *Psychological Science*, 14(3), 262–266. <https://doi.org/10.1111/1467-9280.03432>
- Schmidt, J. R. (2021). Incidental learning of simple stimulus-response associations: A review of color-word contingency learning research. *L'Année Psychologique*, 121(2), 77–127. <https://doi.org/10.3917/anpsy1.212.0077>
- Schmidt, J. R., Crump, M. J. C., Cheesman, J., & Besner, D. (2007). Contingency learning without awareness: Evidence for implicit control. *Consciousness and Cognition*, 16(2), 421–435. <https://doi.org/10.1016/j.concog.2006.06.010>
- Schmidt, J. R., & De Houwer, J. (2012a). Adding the goal to learn strengthens learning in an unintentional learning task. *Psychonomic Bulletin & Review*, 19(4), 723–728. <https://doi.org/10.3758/s13423-012-0255-5>
- Schmidt, J. R., & De Houwer, J. (2012b). Contingency learning with evaluative stimuli. *Experimental Psychology*, 59(4), 175–182. <https://doi.org/10.1027/1618-3169/a000141>
- Schmidt, J. R., & De Houwer, J. (2012c). Learning, awareness, and instruction: Subjective contingency awareness does matter in the colour-word contingency learning paradigm. *Consciousness and Cognition*, 21(4), 1754–1768. <https://doi.org/10.1016/j.concog.2012.10.006>
- Schmidt, J. R., & De Houwer, J. (2019). Cue competition and incidental learning: No blocking or overshadowing in the colour-word contingency learning procedure without instructions to learn. *Collabra: Psychology*, 5(1), 15. <https://doi.org/10.1525/collabra.236>
- Schmidt, J. R., Iorio, C., & Poulin-Charronnat, B. (2023). Automatizing sight reading: Contingency proportion and task relevance in the music contingency learning procedure. *Collabra: Psychology*, 9, Article89743.
- Schultz, B. G., Stevens, C. J., Keller, P. E., & Tillmann, B. (2013). The implicit learning of metrical and nonmetrical temporal patterns. *Quarterly Journal of Experimental Psychology*, 66(2), 360–380. <https://doi.org/10.1080/17470218.2012.712146>
- Schulze, K., Gaab, N., & Schlaug, G. (2009). Perceiving pitch absolutely: Comparing absolute and relative pitch possessors in a pitch memory task. *BMC Neuroscience*, 10(1), 106. <https://doi.org/10.1186/1471-2202-10-106>
- Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42(4), 1096–1104. <https://doi.org/10.3758/BRM.42.4.1096>
- Stoet, G. (2016). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44(1), 24–31. <https://doi.org/10.1177/0098628316677643>
- Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin*, 113(2), 345. <https://doi.org/10.1037/0033-2909.113.2.345>
- Theusch, E., & Gitschier, J. (2011). AbsolutePitch twin study and segregation analysis. *Twin Research and Human Genetics*, 14(2), 173–178. <https://doi.org/10.1375/twin.14.2.173>

- Tillmann, B., & Poulin-Charronnat, B. (2010). Auditory expectations for newly acquired structures. *Quarterly Journal of Experimental Psychology*, *63*(8), 1646–1664. <https://doi.org/10.1080/17470210903511228>
- Tillmann, B., Stevens, C., & Keller, P. E. (2011). Learning of timing patterns and the development of temporal expectations. *Psychological Research*, *75*(3), 243–258. <https://doi.org/10.1007/s00426-010-0302-7>
- Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*(4), 552–564. <https://doi.org/10.1037/0096-3445.134.4.552>
- Vachon, F., Marois, A., Lévesque-Dion, M., Legendre, M., & Saint-Aubin, J. (2018). Can ‘Hebb’ be distracted? Testing the susceptibility of sequence learning to auditory distraction. *Journal of Cognition*, *2*(1), 1–16. <https://doi.org/10.5334/joc.8>
- Van Hedger, S. C., Heald, S. L. M., & Nusbaum, H. C. (2019). Absolute pitch can be learned by some adults. *PLOS ONE*, *14*(9), e0223047. <https://doi.org/10.1371/journal.pone.0223047>
- Ward, W. D. (1999). Absolute Pitch. In D. Deutsch (Ed.), *The Psychology of Music* (2nd ed., pp. 265–298). Academic Press. <https://doi.org/10.1016/B978-012213564-4/50009-3>
- Werker, J. F., & Tees, R. C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *The Journal of the Acoustical Society of America*, *75*(6), 1866–1878. <https://doi.org/10.1121/1.390988>
- Wong, Y. K., Lui, K. F. H., Yip, K. H. M., & Wong, A. C.-N. (2020). Is it impossible to acquire absolute pitch in adulthood? *Attention, Perception, & Psychophysics*, *82*(3), 1407–1430. <https://doi.org/10.3758/s13414-019-01869-3>
- Wong, Y. K., Ngan, V. S., Cheung, L. Y., & Wong, A. C.-N. (2020). Absolute pitch learning in adults speaking non-tonal languages. *Quarterly Journal of Experimental Psychology*, *73*(11), 1908–1920. <https://doi.org/10.1177/1747021820935776>
- Woods, K. J. P., & McDermott, J. H. (2018). Schema learning for the cocktail party problem. *Proceedings of the National Academy of Sciences*, *115*(14). <https://doi.org/10.1073/pnas.1801614115>
- Wulf, G., Höß, M., & Prinz, W. (1998). Instructions for motor learning: Differential effects of internal versus external focus of attention. *Journal of Motor Behavior*, *30*, 169–179. <https://doi.org/10.1080/00222899809601334>

Supplementary Materials

Peer Review Communication

Download: https://collabra.scholasticahq.com/article/118806-incidentally-acquiring-pitch-label-associations-with-a-musical-contingency-learning-task/attachment/230441.docx?auth_token=mhOiv6yt2bCGpmQFIZVp
