THESE DE DOCTORAT DE L'ETABLISSEMENT UNIVERSITE BOURGOGNE FRANCHE-COMTE

PREPAREE AU
Laboratoire d'Étude de l'Apprentissage et du Développement - CNRS UMR5022

Ecole doctorale n°554

Environnement-Santé

Doctorat de Psychologie Cognitive

Par

Claudia Iorio

Acquiring music information: An incidental learning approach.

Thèse présentée et soutenue à Dijon, le 09/12/2022

Composition du Jury :

Prof, De Houwer, Jan	Professeur, Ghent University	Rapporteur
Prof, Rohrmeier, Martin	Professeur, École Polytechnique Fédéral de Lausanne	Rapporteur
DR, Poulin-Charronnat, Bénédicte	Directrice de Recherche, Université de Bourgogne	Examinatrice
DR, Németh, Dezso	Directeur de Recherche, Université Lyon 2	Examinateur
Prof, Lemercier, Céline	Professeur, Université Toulouse Jean Jaurès	Examinatrice
Prof., Schmidt, James	Professeur, Université de Bourgogne	Directeur de thèse

ÉCOLE DOCTORALE
PRES I Bourgogne I Franche-Comté
Environnements - Santé

Titre : Apprentissage d’informations musicales : une approche d’apprentissage contingente.

Mots clés : Apprentissage musicale, apprentissage incidente, apprentissage contingente, lecture a vu, identification des tons, perception du temps.

Abstract: Cette thèse contient mes travaux empiriques de trois années d’étude de l’apprentissage contingent, c’est-à-dire la capacité humaine à apprendre les régularités entre deux ou plusieurs événements, appliquée à la musique. Apprendre la musique demande du temps et des efforts. Cependant, de nombreuses compétences peuvent être automatisées de manière moins chronophage et moins laborieuse. En effet, certaines recherches suggèrent que de nombreuses informations musicales sont, pour la plupart, acquises implicitement. Dans le Chapitre 1, l’avantage potentiel de l’utilisation d’une procédure d’apprentissage incident pour automatiser les sous-compétences musicales utiles pour la lecture à vue et l’identification de la hauteur est discuté. Dans le Chapitre 2, une première série d’expériences étudie si une tâche d’apprentissage contingent peut être utilisée pour apprendre facilement et rapidement à des nonmusiciens les associations entre la position d’une note sur la portée musicale et son nom (c’est-à-dire la lecture à vue). Comme supposé, des améliorations robustes des performances de lecture à vue ont été observées. Dans le Chapitre 3, une deuxième série d’expériences a d’abord exploré si une procédure d’apprentissage contingent peut aider des nonmusiciens à acquérir des associations entre la hauteur et le nom d’une note (i.e., identification des tons). En lien avec notre hypothèse, un apprentissage robuste a été observé. Dans un second temps, une étude a également mis en évidence une amélioration des performances d'identification des tons chez des musiciens. Enfin, les expériences du Chapitre 4 ont étudié la relation entre l'apprentissage contingent et la perception du temps dans l'acquisition des associations entre la hauteur et le nom d'une note. En particulier, ces expériences ont étudié si la prévisibilité temporelle aide à acquérir les contingences par rapport à l'imprévisibilité temporelle. Les résultats de ces études suggèrent que l'apprentissage contingent n'est pas fortement influencé par la perception du temps. Le Chapitre 5 discute plus largement des résultats de cette thèse, en particulier sur la façon dont ouvrent de nouvelles perspectives pratiques d'apprentissage incident pour l'éducation musicale. La relation entre la perception du temps et l'apprentissage contingent est discutée plus en détail.

ÉCOLE DOCTORALE
PRES | Bourgogne | Franche-Comté
Environnements - Santé

Title: Acquiring music information: An incidental learning approach.

Keywords: Music learning, incidental learning, contingency learning, sight-reading, pitch identification, time perception.

Abstract: This thesis contains my empirical works resulting from three years of studying contingency learning, that is the human ability to learn regularities between two or more events, applied to music. Learning music requires time and effort. However, many skills can be automatized in less time-consuming and effortful ways. Indeed, some research suggests that many elements of music knowledge are mostly implicitly acquired. In Chapter 1, the potential benefit of using an incidental learning procedure to automatize musical subskills useful for sight-reading and for pitch identification is discussed. In Chapter 2, the first set of experiments investigate whether an incidental contingency learning task can be used to easily and rapidly teach nonmusicians the associations between note positions on the musical staff and their note names (i.e., sight-reading). As hypothesized, robust improvements in sight-reading performance were observed. In Chapter 3, a second set of experiments first explored whether an incidental learning procedure can be used to help nonmusicians to acquire pitch-label associations. As hypothesized, robust learning was observed. Improvements in pitch identification performance were also observed for musicians. Finally, the experiments in Chapter 4 studied the relationship between incidental learning and time perception in acquiring pitch-label associations. In particular, the experiments investigated whether temporal predictability aids in the acquisition of contingencies relative to temporal unpredictability. Extending previous research, the evidence gathered from these studies suggested that contingency learning is not strongly influenced by time perception. Chapter 5 discusses the results of this thesis more broadly, with particular focus on how the present results open up new venues for practical implementations of incidental learning procedures in the music education domain. The relationship between time perception and contingency learning is further discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank Jan De Houwer, Martin Rohrmeier, Céline Lemercier, Dezso Németh, and Bénédicte Poulin-Charronnat, for having accepted to be part of my jury.

A huge thanks to James R. Schmidt, for all the valuable comments and feedback. Despite all the lockdowns that obliged us to work from home, I never felt alone on this journey.

I would like to deeply thank the LEAD members for creating an enjoyable working environment. A special thanks to Bénédicte for always being there whenever I needed help. Lots of thanks to Corinne and Sandrine for helping even when they were not supposed to. Thanks to all my colleagues, in particular Iva, Ella, and Yannick for the nice moments spent together in office 215.

Thank Sonja Kotz for having welcomed me during my research stay at the BandLab in Maastricht and thank you to all the colleagues there for the enjoyable moments together.

Mille, cento ed infiniti grazie alla mia famiglia, mamma, papà e Giulia per avermi sempre supportato ed aver accetato le mie scelte. Siete la mia forza e sempre con me ovunque io sia. E Andrea, un fratello acquisito, che con gli anni ho imparato a conoscere ed apprezzare, grazie per i momenti spensierati e le risate.

Last but not least, thanks to Damien, you are the most wonderfully unexpected surprise I found in Dijon. Thank you for being there no matter what.

Thesis valorization

Publications

- **Iorio C.**, Šaban I., Poulin-Charronnat B., Schmidt J.R., "*Incidental Learning in Music Reading: The Music Contingency Learning Task*", (2022, Quarterly Journal of Experimental Psychology).
- Schmidt J.R., **Iorio, C.**, Poulin-Charronnat B., "*Automatizing sight-reading: Contingency proportion and the task relevance in the musical contingency learning procedure*", (Submitted).
- **Iorio, C.**, Schimdt, J.R., "*Incidentally acquiring pitch-label associations with a musical contingency learning task*", (in preparation).
- **Iorio, C.**, Schimdt, J.R., Poulin-Charronnat, B., and Kotz, S., *"Learning music through time perception"*, (in preparation).

Communications

2022: "Experimentarium", communications about my thesis addressed to scholars and nonscientific community (scientific vulgarization).

2021: Iorio, C., Šaban, I., Poulin-Charronnat, B., Schmidt, J., R. "Acquiring Musical Skills Without the Goal to Learn", at the Colloque des Jeunes Chercheurs/es, online conference organized by Fresco.

2022: Iorio, C. & Schmidt, J., "Incidentally acquiring pitch-label associations with a musical contingency learning task", Escop, Lille FR.

2022: Iorio, C., Witt A., Schmidt, J., Tillmann, B., Lagarrigue, Y., "The effect of auditory rhythmic priming on motor sequence learning"**,** at the Forum des Jeunes Chercheurs/es, Dijon FR.

2021: Iorio, C., Šaban, I., Poulin-Charronnat, B., Schmidt, J., R. "Incidental learning in music reading: a modified version of the musical Stroop task", at the ICMPC, online.

2021: Iorio, C., Šaban, I., Poulin-Charronnat, B., Schmidt, J., R "Incidental music learning with a contingency paradigm", at the Forum des Jeunes Chercheurs/es, online.

Chapter 1 - GENERAL INTRODUCTION

Many people, at least once in life, try to learn how to play a musical instrument. Where I grew up, music education at school does not really encourage people to learn any instruments. The common idea is that to be able to play you must have a “gift”, something innate that allows you to understand music and be able to play an instrument. So, as soon as I started to study Viola at the Conservatory, I was considered as a gifted person lucky enough to be able to magically understand music and play the viola. However, the true story is a little bit different and, contrarily to what people often believe, learning how to play implied motivation, effort, and above all practice. This is at least what I did; for ten years and more, I studied, I motivated myself to improve, and above all I spent countless hours practicing. About three years ago, I read a PhD offer that aimed to study the role of incidental learning (learning without the goal to learn, the opposite of my music education) in acquiring musical information. Given my experience in learning music, I was curious about this subject, so I applied and fortunately I had the chance to start working on the topic. Since then, I worked on investigating how an incidental procedure might aid in the acquisition of musical information useful to learn how to play. Here is a little example of implicit or incidental learning from real-life: It happens to many of us to find ourselves starting to sing the song from the supermarket where we go for groceries. Now, I doubt that any one of us would like to intentionally learn the music of the supermarket; but because we are exposed to it many times, we learn the jingle. This is just a real-life example of incidental learning. However, based on more systematic explorations, science suggests that music information is mostly acquired implicitly through mere exposure (Bigand & Poulin-Charronnat, 2006; Ettlinger et al., 2011; Rohrmeier & Rebuschat, 2012). During my thesis, I explored some questions related to the process of implicitly acquiring music information. In particular, this work focused on incidental learning in music, that is the ability to learn something without the intention to do it. The big question I tried to answer is whether it is possible to incidentally learn musical information. My research here mostly focuses of connecting cognitive psychology and music cognition in the attempt to investigate the potential role of incidental learning in acquiring musical information important to learn how to play an instrument.

The following pages are divided into different chapters. In the first chapter, I present the theoretical background where I talk about music cognition in a general way and more specifically about the process of music learning. I begin by discussing the important role of practice in automatizing complex skills, for instance, how to play an instrument. Then, I introduce the research on one particular way to incidentally acquire new information, namely with a contingency learning procedure. Finally, I end this section with a discussion about how learning can be influenced by time perception. Specifically, I ask the question: Does our perception of time influence the way we process information and thus how we learn about these stimuli?

In the second, third, and fourth chapters of the thesis I present the experimental works developed during the three years. The studies are grouped into three main topics: sight-reading (Chapter 2), pitch identification (Chapter 3), and time perception (Chapter 4). All the three topics are discussed with regard to the process of incidentally acquiring new information.

The last chapter of the thesis is a general discussion on the contribution of my results to the research presented in the theoretical part of the thesis. I further present some openings for future possible works on this line of research as well as potential practical uses in real-life music education. Finally, a brief section is dedicated to the side research projects that I conducted during the three years of thesis, which have some meaningful relationship to the thesis work.

Music cognition

Music is a universal and complex activity that involves many different cognitive processes (Honing & Ploeger, 2012; Pearce & Rohrmeier, 2012). Because of its complexity, music has been at the center of scientific research for long time with regards, for instance, to the way that human beings perceive music, produce it, and learn it (Pearce & Rohrmeier, 2012, for a review). Research has been conducted from both experimental and computational perspectives, for instance, using a finite-state grammar procedure Loui et al. (2010) demonstrated that both adults and infants are able to learn long melodic sequences. On the other side research implementing neural networks and computational models focused on the investigation of music production, retrieval, and perception (Rohrmeier & Rebuschat, 2012, for a review). In the current thesis, I am particularly interested in a specific aspect investigated from an experimental perspective, not only for purely theoretical reasons but also for the potential practical implications and this is the central topic of which I am going to speak about in this thesis: the process of learning music.

Most people tend to think that music is, exclusively or predominantly, learned in an explicit way. That is, if you want to become a musician you need to deliberately/intentionally spend a lot of hours practicing on your instrument. Globally this idea is true, and much research has linked deliberate learning with musical expertise (Ericsson et al., 1993; Ericsson & Harwell, 2019a; Lehmann, 1997, see Macnamara 2014 for a related issue in nonmusical domain, such as sports). Specifically, although there is not always consensus on the definition of deliberate practice, some authors have pointed out that to achieve expertise in music a certain amount of time is required, where the students intentionally implements conscious strategies to improve performance (see How et al., 2021, for a review on the different practice strategies implemented by students).

However, a large body of research also suggested that music can be implicitly acquired, with regards, for instance, to the implicit acquisition of melodies (e.g., Saffran et al., 1999, 2000; Tillmann & Poulin-Charronnat, 2010), harmony (e.g., Bly et al., 2009; Loui et al., 2010; Rohrmeier & Cross, 2009), timbre (e.g., Bigand et al., 1998), and temporal sequences (e.g., Brandon et al., 2012; Salidis, 2001; Schultz et al., 2013; Tillmann et al., 2011). For instance, Saffran et al. (2000) in a series of studies, tested whether infants were able to recognize familiar music to which they were previously exposed. Their results suggested that although the infants were merely exposed to music, they were able to learn the musical extracts showing a preference for the familiar music compared to the unfamiliar one.

As another example, Schultz et al. (2013) demonstrated that participants involved in a Serial Reaction Time Task (SRTT), were able to implicitly learn auditory temporal patterns. That is, participants listened to auditory stimuli that, unbeknownst to the participants, followed three different metrical strengths (the stimuli could be either Strongly Metrical, Weakly Metrical, or Nonmetrical based on when the stimulus aligns with event onsets at the level of the pulse).

In sum, these prior results suggested that music information can be implicitly acquired, however, most of this past research focused on perceptual processes in listeners. That is, interest was particularly focused on the investigation of how listeners process musical information they listen to. Here, the aim of the present work is different from previous research about music and implicit or incidental acquisition primarily because I focus on music performance and more precisely on the acquisition of musical information useful to learn how to play.

Music performance and automaticity

Music performance (such as playing an instrument) refers to a variety of different tasks. For instance, Sloboda (2000) distinguished between a technical component and an expressive one. The first relates to a series of mechanical subskills that produce “fluent coordinated output”, for instance the movement of the fingers on a piano board. The second refers to “intentional variations in performance parameters chosen by the performer to influence cognitive and aesthetic outcomes for the listener”, for instance whether a performer decide to press the bow on the violin strings with more or less strength, which will produce a change in the production of the sounds that in turn will influence the way that the listener perceives that sound (e.g., it can be a “piano” if the pressure of the bow on the string is weak or a “forte” if the bow is pressured more).

One of the biggest differences between skilled and novice musicians is the automatization of subskills useful to play. If you have ever tried to learn to play the violin, or any other instruments, you can easily relate to this process: At first, it seems almost impossible to control everything, for instance the bow movements, the left fingers on the violin board, the reading of the notes, and the production of the accurate motoric response. However, with enough practice you can actually “feel” that some of these subskills becomes less demanding and more automatic. For example, you will have the feeling that your brain can read and decode the musical notes on your sheet music without any effort (e.g., as easy as it is to read a book).

Of course, learning how to play an instrument requires time and effort, similarly to acquiring any other new skill, such as playing chess or speaking a new language. Skill acquisition usually follows three phases (Anderson, 1982; Fitts & Posner, 1967; Schneider & Shiffrin, 1977). At first, the task is new, and it is characterized by controlled cognitive processes. That is, we engage in intentional and conscious strategies to achieve the instructed task. The second phase of the process involves a combination between controlled and automatic processes. Control is still being exerted to master the task, but through experience the appropriate actions to the stimuli presented to us become gradually automatized. For example, a stimulus triggers the appropriate response in an unconscious and effortless manner. In the last step, the skill has been acquired and it is completely automatic, that is, cognitive control is no longer required. Therefore, the skill acquisition process is characterized by a transition from a more initial and controlled stage to a more automatic and less cognitive demanding phase. How is it possible to automatize skills? One answer to this question is practice (Langan-Fox et al., 2002; Logan, 1988). Whether you want to be fluent in Italian, play football like Maradona, or play the cello like Yo-Yo-Ma you need to repeat the same task many times.

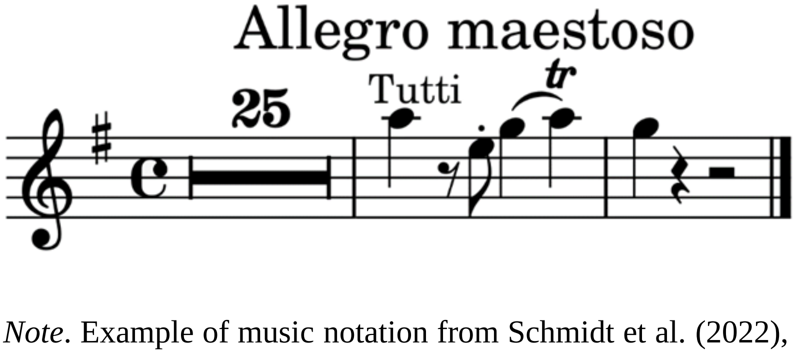
Here, I specifically investigated the possible utility of an incidental learning procedure in benefiting the automatization of subskills useful for music performance. However, it is not the aim of this work to discuss the process of music performance in all its complexity and subcomponents, because this would represent a complex and very long research project that would certainly require more than the time I had for my thesis. Therefore, here I am going to focus on the acquisition of some of the specific musical subskills needed for the technical component of music performance: sight-reading and pitch identification.

Sight-reading and automaticity

Sight-reading is the ability to read a musical piece written in standard musical notation (i.e., the symbols used to visually represent music, which consists of a musical staff on which the notes are written with note forms to indicate durations and vertical note placement to indicate the pitch) and perform it on your instrument without previously seeing or practicing the novel piece (Waters et al., 1997; Wolf, 1976). Sight-reading involves many different factors, such as the coding of visual information, the preparation of motor responses, and visuomotor integration (Kopiez & In Lee, 2006, 2008; Lehmann & Kopiez, 2009; Gudmundsdottir, 2010). It is usually taught in an explicit way (Ericsson et al., 1993; Ericsson

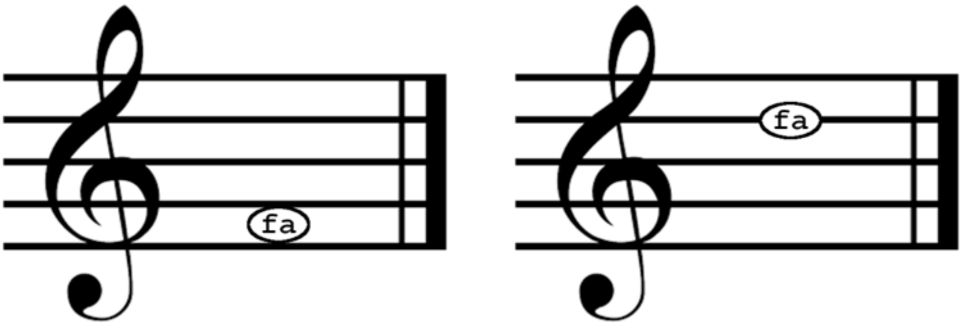
& Harwell, 2019; Hébert & Cuddy, 2006; Lehmann, 1997; Mills & McPherson, 2006; Mishra, 2014). For example, in aural training, musicians are tasked with melodic dictation to improve their ability to recognize pitches. In solflege/singing, they are required to name and to sing out loud the name of the notes presented on the musical sheet. This is widely used to teach sight-reading (Mishra, 2014) . Strangely, although students spend many years looking at sheet music and playing this written musical information on their instrument, sight-reading skills are usually lacking among music students (Hargreaves, 1986; Mills & McPherson, 2006; Scripp, 1995), often even rather advanced ones.

If we ask someone without any music theory knowledge to read a musical sheet, such as the example presented in Figure 1, they will inevitably find it almost impossible to understand the meaning of the symbols on the music staff. This is because someone that did not study music does not possess a meaningful knowledge of the "translation" of music notation. To them, the symbols on the music staff are just drawings without any meaning, a situation that is similar to someone trying to read Chinese even when they have never studied it or for a baby without any reading skill to try to read in any language (see Catale & Meulemans, 2009, Montgomery & Koeltzow, 2010, for instance).

Figure 1- Example of music notation.

Note. Example of music notation from Schmidt et al. (2022), which represents a flute excerpt from the "Flute Concerto No. 1 in G major, K. 313" by W. A. Mozart.

On the other hand, musicians can easily and automatically read music, as indicated, for example, by research using Music Stroop procedures (e.g., Crump et al., 2012; Drost et al., 2005; Grégoire et al., 2013; Stewart, 2005; Zakay & Glicksohn, 1985 for musical Stroop procedures). Similar to the classic Stroop procedure (MacLeod, 1991; MacLeod & MacDonald, 2000, for nonmusical Stroop procedure), in musical Stroop tasks, participants are asked to identify the note name (often by reading it aloud) written inside of a note presented in either a congruent position (vertical location) on the music staff (e.g., "fa" written inside of the note for "fa") or in an incongruent position (e.g., "fa" written inside the note for "re"), as shown in Figure 2. The results from these types of tasks showed a significant difference between congruent and incongruent trials in Reaction Times (RT) for musicians, but not for nonmusicians. Specifically, musicians responded much faster to congruent than to incongruent trials, suggesting that they automatically processed the note-position, and this had an impact on note-name reading. In other words, they automatically "read" the note position, which can aid in responding when the note name matches the position or slow down responding when the note name indicates a different, conflicting response. While this effect, typically referred to as the Musical Stroop Effect (MSE), is systematically found in musicians, unsurprisingly nonmusicians did not show the effect, suggesting that nonmusicians can easily read the note name without any influence from the note position. Nonmusicians are not influenced by the incongruency between the note name and the note position, because they do not know the meaning of the musical symbols and therefore, they cannot translate the note position into a meaningful outcome.

Figure 2 - Example of congruent and incongruent trials.



Note. Example of congruent trial on the left in which the note position for “fa” matches with the name written inside the note. On the right is an example of incognruent trial, in which the note position for “ré” mismatches with the note name.

For decades, the (nonmusical) Stroop effect (Augustinova & Ferrand, 2014; MacLeod, 1991) has been used to describe the automatic influence of task-irrelevant stimuli (e.g., word) on performance of another task (e.g., naming the color). That is, in the classic color-word Stroop effect, although participants are instructed to name the color, they cannot avoid reading the word, resulting in increased RTs when the color and the word mismatch compared to when they match. This result suggested that reading is an automatic process, in other words, seeing the word is sufficient to provoke a bias to give the corresponding response although this is not what it is required in the task. In the current work, although I am going to use a Stroop-like procedure, the aim is to show that a similar MSE can be found in nonmusicians after a short acquisition phase using an incidental learning procedure. The presence of the MSE in nonmusicians would suggest that a complex subskill such as sight-reading can be easily and quickly automatized thanks to the rapid acquisition of the note name/note position associations through the incidental learning task.

Pitch identification and automaticity: the case of Absolute Pitch

In a similar series of experiments, I will investigate the possible utility of an incidental learning procedure to aid in learning to detect pitches by ear (e.g., to hear a note and say whether it is do, re, mi, etc.). Pitch identification is quite unique in music. For instance, while the majority of people are able to recognize the color blue just looking at it and without any need to compare the blue with another color like yellow, in music most of the musicians rely on a Relative Pitch (RP) strategy when identifying pitches. Specifically, most musicians are able to identify and name pitches based on a comparison strategy (Levitin, 1994; Levitin & Rogers, 2005; Takeuchi & Hulse, 1993) in which they must already know the identity of one reference tone so that they can compare the interval (i.e., numbers of the semitones, akin the smallest distance between notes) between the tone they are listening to (e.g., “mi”) with the known reference tone (e.g., “do”, in this case specific the interval between the two tones is a major third, meaning that the “mi” is four semitones above the “do”, see Figure 3). In other words, most people (including the majority of skilled musicians) cannot simply listen to a tone out of context and identify it by name (termed Absolute Pitch). However, if we play to RP possessors a reference tone first and tell them which note it is, then they may be able to determine the identity of a second note that we play for them.

Figure 3 - Example of semitones.

Note. The picture represents a piano keyboard, the semitones are the distance between each key, for instance the interval between the note “do” and “mi”, called a major third, in which “do” is four keys or “semitones” away from “mi”.

RP possessors are characterized by less automatization of the pitch identification process, as suggested by their slow RT in pitch identification tasks (Bermudez & Zatorre, 2009; Miyazaki, 1990; Takeuchi & Hulse, 1993; Van Hedger et al., 2019; Wong, Lui, et al., 2020). That is, even when they can identify a pitch using an RP strategy, this identification takes time. Absolute Pitch (AP), on the other hand, is a much faster strategy that seems to rely on a more automatic retrieval (i.e., retrieving the pitch name directly from the long-term memory).

In other work, procedures that implemented an auditory version of the Stroop task (Akiva-Kabiri & Henik, 2012), showed an asymmetrical effect for AP possessors compared to non AP possessors. That is, people with AP are more impaired when asked to identify the tones compared to people that rely on RP strategy. This is not surprising when considering that only AP possessors seem to rely on an automatic retrieval of the pitch-label association from memory.

Most of the existing research has argued that AP may be due to an interaction between individual differences, such as age of beginning musical training and genetic components (Athos et al., 2007; Baharloo et al., 2000; Crozier, 1997; Deutsch, 2013a, 2013b; Deutsch et al., 2006; Miyazaki & Ogawa, 2006; Theusch & Gitschier, 2011). For instance, some research (see Deutsch, 2013a, for a review), shows that there is a higher percentage of AP possessors in musicians that started musical training before the age of 4 compared to musicians that started musical training later. Furthermore, Theusch and Gitschier (2011), reported higher percentage of AP possessors in monozygotic twins compared to heterozygotic twins, suggesting that genetics can play a role in the ability to learn AP information.

However, while genetics and musical training may indeed play some role in how easy it is for a given person to learn AP, the “learning theory” hypothesizes that AP can be acquired throughout lifespan by almost anyone (i.e., even without the “good” genetics or early musical training). For instance, a few investigations (Van Hedger et al., 2019; Wong, Lui, et al., 2020; Wong, Ngan, et al., 2020) have shown that with an effortful and explicit training it is possible to achieve AP to a strict criterion. On the other hand, such results are sometimes questioned as the very small number of participants that reach the AP criterion typically have pretest AP scores that are already close to said criterion.

Although it is not the aim of this work to claim that it is possible to learn AP using an incidental learning procedure (perhaps a goal for future longer-term training studies), I want to study whether it is possible to incidentally strengthen and automatize the link between the pitch and its label to help improving pitch identification performance. Therefore, in the second section of this thesis I will use an auditory Stroop-like procedure to study the role of an incidental task on the acquisition and identification of pitch-label associations.

Incidental learning: the case of contingency learning

It is well known that human beings can implicitly or incidentally acquire new information. For instance, after listening to a few minutes of pseudowords that are created with artificial grammar rules (e.g., which letters can or cannot follow which), we are able to identify with reasonable accuracy which new pseudowords are consistent with the grammar and which are not (Reber, 1967; for a review, see Pothos, 2007). This artificial grammar learning occurs even though learning is incidental, meaning that participants are not informed that there *is* a grammar to learn (e.g., they are often falsely told that the goal is to memorize the specific stimuli presented), and participants are often unaware that they have learned anything. Similarly, in a simple stimulus identification task where, unbeknownst to participants, the responses follow a predictable (repeating) pattern, participants will learn the sequence in a similarly implicit way (Nissen & Bullemer, 1987; Turk-Browne et al., 2005). Interestingly, we are able to implicitly (or incidentally) learn new information and use this information in a variety of tasks (e.g., in language acquisition; see Aslin et al., 1998; Saffran, Aslin, et al., 1996;

Saffran et al., 1997; Saffran, Newport, et al., 1996). Last but not least, implicit learning procedures, such as sequence learning (Nissen & Bullemer, 1987; Turk-Browne et al., 2005), 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), produce a rapid learning effect in a manner of minutes from the start of the task. While not everything can (or should) be learned in a purely incidental or implicit way, the speed of this type of learning can be particularly useful for automatizing certain subskills, as illustrated in the current thesis with sight-reading and pitch detection learning.

Before turning to the main topic of my thesis, I would like to introduce a small clarification about the type of learning that I will discuss here. There is quite a debate on the implicit or explicit knowledge acquired through implicit or incidental learning tasks (Cleeremans et al., 1998). Whether participants are trying to learn (deliberate learning) or unaware that there is something to learn (incidental learning) and whether participants do or do not become aware of what they learned are two correlated yet different issues. It is not my aim to discuss the debate about whether learned information is conscious or unconscious any further. Rather, I focus on the potential utility of incidental learning procedures. For this reason, I will use the term “incidental learning” in the rest of this thesis to refer to the acquisition of new information without the goal to learn (Kerka, 2000). The main topic of this thesis is to discuss the incidental acquisition of new material. In particular, I am interested in one particular methodological approach to incidentally learn new information: a contingency learning task.

Contingency learning refers to the human ability to detect regularities between events in the environment, and contingency learning tasks are one way to study rapid incidental learning of the regularities between two or more events (e.g., Event B tends to follow Event A, making Event A a predictive cue for Event B; for reviews, see MacLeod, 2019; Schmidt, 2021;

for related learning procedures, see Carlson & Flowers, 1996; Miller, 1987; Mordkoff & Halterman, 2008; Musen & Squire, 1993). Usually in a contingency learning procedure, participants are asked to respond to a relevant stimulus or target while a regularity between the target (e.g., a color, Schmidt et al., 2007) and an irrelevant stimulus or cue (e.g., a word, Schmidt et al., 2007) is presented. Although participants are not asked to explicitly learn the regularities, exposure to the regularity is enough to learn. They learn the co-occurrence between the target and the cue and they start to respond faster and more accurately to events that are consistent with the learned regularity versus events that are unpredictable or incompatible with it (Perruchet, 2019; Perruchet & Pacton, 2006; Thiessen et al., 2013).

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, 2012d), and in all these studies people were able to learn the regularities presented. The pattern of results is always the same, that is, faster reactions times for the most presented and predictable regularity compared to the infrequently presented events.

This thesis takes particularly inspiration from the color-word contingency learning procedure of Schmidt et al. (2007). Similar to a color-word Stroop procedure, participants are asked to respond to the target (the color), while ignoring the cue (the word). However, a regularity is introduced between these two stimulus dimensions to study associative learning. The neutral words (color-unrelated) are presented much more often in one color (high-contingency trials; e.g., “move” most often in blue, “sent” most often in red, etc.) than in the other colors (low-contingency trials; e.g., “move” in red, “sent” blue, etc.). Although participants are not informed of the contingencies between colors and words and often do not become aware of the manipulation, they respond more quickly and more accurately to *high-contingency* trials than to *low-contingency* trials (Schmidt & De Houwer, 2012b).

The main aim of this thesis is to apply results from the contingency learning research to the music domain. In particular, I am interested in studying whether nonmusicians can easily and rapidly acquire music information via an incidental contingency learning task. As mentioned above, sight-reading and pitch detection are typically considered to be hard to master, either taking very long to perfect (in the case of sight-reading) or being entirely impossible without early acquisition and/or the right genetics (in the case of pitch identification). However, the central hypothesis of this thesis is that both of these skills are learnable and can probably be learned much more quickly than previously assumed if the same principles of rapid learning in the nonmusical domain also apply to learning about musical materials.

Time perception and learning

A final part of this thesis is dedicated to the investigation of the relationship between time perception and learning. Recently, there is an increasing interest to the role of temporal information and its influence in the learning process. For instance, Selchenkova et al. (2014) showed that temporally regular presentation of the stimuli (i.e., when a stimulus is presented on a metrically regular base, for instance, when the intertrial onset interval between the tones is constant) benefits the implicit learning of an artificial pitch grammar relative to temporally irregular presentation (i.e., when the tones are presented in randomly varying intervals; see also Geiser et al., 2012; Lange, 2009; Schmidt-Kassow et al., 2009; Schwartze et al., 2011). The authors suggested that external regularities can help listeners to develop perceptual expectations about the temporal occurrence of future tones, thus facilitating the learning of the pitch grammar. Of particular interest for this report is the role of temporal perception in human contingency learning.

In contingency learning research, the role of time has been investigated in a few reports, especially with regards to the temporal contiguity hypothesis (Buehner, 2005). According to this perspective, the acquisition of the contingencies is due to the closeness in time between the presented stimuli. For example, it is easier to notice that two stimuli tend to be presented one after another if the second stimulus is presented very shortly after the first. On the contrary, Schmidt and De Houwer (2012) proposed that the contingency learning effect is not directly influenced by temporal contiguity. Their results support a temporal insensitivity hypothesis, suggesting that for a performance task (when an explicit judgment is not required) temporal contiguity does not seem to notably influence the size of the acquisition of the contingencies. Specifically, when words were presented before the target color, increasing the interval of time between the word and color presentation did not seem to meaningfully impact the size of the learning effect. However, according to the temporal coding hypothesis, temporal information (the "when") is encoded at the same time as the identity of the events (the "what") (Greville & Buehner, 2010; Matzel et al., 1988), and this may actually have an effect on the acquisition process itself. With this hypothesis, the role of time in associative learning goes beyond the idea of temporal contiguity, suggesting the need to look differently at the temporal structure and try to explain its role in this kind of learning. In particular, Balsam et al. (2010) suggested that learning involves extracting the temporal structure of the events and that this temporal information may influence learning of the contingencies itself. Furthermore, previous results suggested that temporal predictability between events (regular intervals between stimuli) can be used to promote casual inference (Greville & Buehner, 2010) and thus strengthening the association between the cue and the outcome. Following this line of research, these studies aim to investigate the role of temporal predictability in the acquisition of contingencies.

Experimental part and research goals

In the following pages I present my experimental work, the studies I did during my thesis as empirical support to my research. The studies are presented in their manuscript versions, and they are grouped in three different topics. Overall, the main research goal is to apply results from contingency learning research to the acquisition of musical subskills. Specifically, the main research question is to investigate the ability of incidental learning procedures to benefit the automatization of musical subskills useful for sight-reading and for pitch identification.

The first set of studies target sight-reading skills as the main point of research. Specifically, the experiments focus on investigating whether it is possible to apply incidental contingency learning rules to easily and rapidly teach nonmusicians the associations between a note position on the musical staff and its name. Firstly, I hypothesized that using a contingency learning procedure is possible to promote an effortless and quick acquisition of the note name/note position association. Specifically, the repetitions of the stimulus-response associations can help the acquisition of the visuomotor translation useful in sight-reading performance.

In the second set of studies, the focus of the research is on pitch identification ability. I investigated whether an incidental learning procedure can be used to help nonmusicians to acquire the pitch-label associations and whether this kind of task can produce an improvement of pitch identification in musicians, as well. I also investigated the role of incidental learning as opposed to deliberate learning (i.e., the process of intentionally acquire new information) for memory consolidation of pitch-label associations. In the second series of studies, I hypothesized that, similarly for the visual domain, using an incidental learning procedure it is possible to strengthen note name-pitch associations that in turn can benefit the pitch identification process.

Finally, the last studies are about the relationship between incidental learning and time perception in acquiring pitch-label associations. Specifically, I wanted to investigate the role of time perception in acquiring pitch-label associations using an incidental contingency learning procedure. The main aim is to promote the investigation of the link between time perception and acquisition specifically in the context of contingency learning. However, it is worth to note that the study reported here is an initial work on this line of research. As first hypothesis, the idea is that time perception can have an influence on the way we learn pitch-label associations.

Chapter 2 – learning and sight-reading

Incidental Learning in Music Reading: The Music Contingency Learning Task

Claudia Iorio, Iva Šaban, Bénédicte Poulin-Charronnat, and James R. Schmidt

Article published in Quarterly Journal of Experimental Psychology
DOI: 10.1177/17470218221092779

Introduction

Music is a complex ability that involves a range of different cognitive processes (e.g., learning, perception, production; Pearce & Rohrmeier, 2012). Not surprisingly, then, during traditional music instruction a wide range of skills need to be learned, such as familiarization with the instrument and musical theory. While traditional training is well adapted to the acquisition of many of these skills, some skills tend to fall behind. One important musical skill, which takes a considerable amount of time to acquire, is *sight-reading* ability. Sight-reading refers to the ability to look at a new piece of music for the first time and play it while reading (e.g., without having to memorize or practice the piece beforehand). Typically, explicit tutoring and deliberate practice are used to teach and improve sigh-reading abilities (Ericsson et al., 1993; Ericsson & Harwell, 2019; Hébert & Cuddy, 2006; Lehmann, 1997; Mills & McPherson, 2006; Mishra, 2014). However, even after many years of studying sight-reading, these skills are still lacking among many music students (Hargreaves, 1986; Mills & McPherson, 2006; Scripp, 1995). In this paper, we will introduce a novel approach to aiding with sight-reading training, intended as a potential supplement to traditional music instruction. As will be discussed below, our new approach aims to leverage the benefits of incidental learning procedures (e.g., very rapid learning), rather than deliberate practice, to facilitate learning. We note in advance that the current research focuses on one component of sight-reading, namely, responding to the note position stimuli with the corresponding actions.

The difficulty of sight-reading

Part of the difficulty in learning to sight read may be due to the complexity of the task. Indeed, sight-reading is a complex skill that relies on different factors (Kopiez & In Lee, 2006, 2008; Lehmann & Kopiez, 2009) and it involves different processes based on the coding of visual information, motor responses, and visuomotor integration (Gudmundsdottir, 2010). Although the terms "music reading" and "sight-reading" are often used interchangeably, the first can be considered as a prerequisite of the second. That is, while music reading mostly refers to the act of reading and decoding musical notation from music sheets, sight-reading refers to the complex skill that involves different components such as reading and decoding musical notation (i.e., music reading) and performing (playing) the music directly while reading, that is, without prior practice (Waters et al., 1997; Wolf, 1976). Therefore, it has been defined as a demanding transcription task (Sloboda, 1982, 1985).

Schön et al. (2001, 2002) hypothesized that at least three types of translations are involved when musicians read music: singing-like (visual to auditory transcoding), playing-like (visual to motor transcoding), and note-naming-like (visual to verbal transcoding). Accordingly, Stewart et al. (2003) proposed that musicians automatically generate a sensorimotor translation of a spatial code (written music) into a series of motor responses (keypresses). Reading music requires analyzing visual information. In particular, it is necessary to decode the spatial position of the notes on the music staff. While the horizontal location carries information about the duration, the vertical location indicates the pitch (Sergent et al., 1992). Previous research suggested that timing and pitch information (i.e., the horizontal and the vertical positions of the notes on the staff) are perceived and coded separately (Schön et al., 2001, 2002; Stanzione et al., 1990). Here, we focused on the encoding of the vertical position of the notes, a process that has been investigated in some prior research. Sloboda (1976), for instance, compared the performance in a recall task between musicians and nonmusicians. His results showed that nonmusicians were less accurate in recalling a sequence of notes than musicians, suggesting that naming the visual stimulus can be the first step to encode visual material. Perea et al. (2013) further provided evidence that coding the position of the notes relies on more than just visualization. They used a same/different task, in which participants were asked to judge the similarity between two musical sequences. Nonmusicians had worse performance compared to musicians, suggesting that note position coding is quite approximate at early stages of processing compared to more experienced readers.

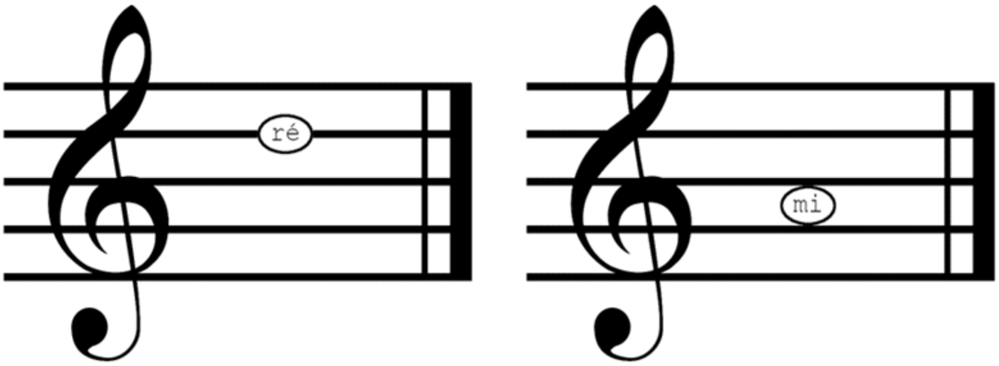
In addition to being a complex task, focal study of sight-reading skills is atypical (Hardy, 1998). Instead, a music practice often involves a focus on mastering music scores, frequently with blocked repetition (Barry, 1992, 2007; Maynard, 2006; Rohwer & Polk, 2006), and a music education often focuses on music theory, instrument technique, etc. These are all important skills as well, of course, but sight-reading, though a valued skill, is often ignored. One difficulty in teaching sight-reading is that students need to automatize the translation of the notes from the page to the actions on the instrument, and for this an enormous amount of novel materials (e.g., music scores) would be needed (Hardy, 1998). For instance, a familiar musical score that the student has already seen and played before is not very useful in practicing the skill of seeing new, unfamiliar material and rapidly playing it while reading.

Automaticity and the Musical Stroop

Though complex, many musicians will eventually automatize their sight-reading skills. Automatizing particular components of a skill is likely to be crucial to learning complex skills and it is often the key for acquiring expertise. For instance, expert chess players are incredible good at reading the board positions, mostly because they can easily and automatically retrieve encoded positions of the chess pieces on the board after years of looking at chessboard configurations (e.g., Saariluoma, 1994).

Similarly, musicians can easily and automatically read music notation. A number of studies using *musical Stroop procedures* (Grégoire et al. 2013; see also, Crump et al., 2012; Drost et al., 2005; Stewart, 2005; Zakay & Glicksohn, 1985, for other musical Stroop procedures), comparing performance between musicians and nonmusicians, provided evidence to support the view of music reading being an automatic process for musicians. Some authors (Grégoire et al., 2013, 2014b, 2014a, 2015, 2019) proposed that this automaticity in musicians may be due by the learned associations between note-positions and note-names in musicians. In musical Stroop tasks, participants are presented with a note on the musical staff with a note-name written inside of it, as illustrated in Figure 4. On congruent trials, the meaning of the note-position (task irrelevant) and the note-name (task relevant) match (e.g., “ré” written inside of the note for “ré”). On incongruent trials, the meaning of the note-position and note-name mismatch (e.g., “mi” written inside the note for “la”). Analogous to color-word Stroop tasks (see MacLeod, 1991; MacLeod & MacDonald, 2000, for nonmusical Stroop procedures), musical Stroop procedures measure the automatic influences of previously learned associations between note positions and their note names on reading simple written note names. Although the task was to ignore the note-position (i.e., where the note was presented on the musical staff) and simply respond to the note-name written inside of it, musicians processed the note-position and this had an impact on note-name reading, as indicated by slower and less accurate responses to incongruent trials relative to congruent trials. This phenomenon has been termed the Musical Stroop Effect. Contrary to the Musical Stroop Effect observed in musicians, nonmusicians responded just as quickly to incongruent as to congruent name-note pairs (i.e., no Musical Stroop Effect). This is unsurprising, as nonmusicians have not learned the meaning (or “translation”) of the note positions (i.e., the association between the note-position and note-name) in the first place and are simply reading the written note names (without any possible influence of the note positions).

Figure 4 - Example stimuli in the musical contingency learning task.



Note. On the left, a congruent stimulus or high-contingency trials ("ré" printed in the note for "ré"). On the right, an incongruent stimulus or low-contingency trial ("mi" printed in the note for "la").

Previous work with musical Stroop procedures studied the influence of the knowledge acquired before participants entered the laboratory. That is, past work has studied the influence of music knowledge that expert musicians already possessed. Our goal is exactly the opposite: to train nonmusicians to acquire music knowledge that they do not yet possess. Unlike previous research using musical Stroop procedure, here we want to demonstrate that by using an incidental training procedure (discussed shortly) nonmusicians can rapidly acquire such automatic influences of music reading akin to the Musical Stroop Effect previously found in skilled musicians. That is, using an incidental training, nonmusicians should show a Musical Stroop Effect, even after very brief training, supporting the idea of a rapid and incidental acquisition of a complex subskill (i.e., music sight-reading skills). We note that although the term "automaticity" has been used to describe many different features of learning (e.g., the need for awareness, attentional and cognitive resource needs, the stimulus- or goal-driven nature of learning; Moors & De Houwer, 2006), it is certainly not our goal to argue that the learning we observe is automatic in all of these senses. Here, we refer to "automaticity" to describe the "automatic" impact of task-irrelevant note positions on performance of another task (i.e., in the same sense that a color-word produces "automatic" influences on color naming in the traditional

Stroop procedure; Augustinova & Ferrand, 2014). That is, we ask whether it is possible that nonmusicians can rapidly acquire similar automatic influences of sight-reading knowledge on behavior as that observed in the Musical Stroop Effect with musicians that have more extensive musical training.

Incidental contingency learning

Our research applies knowledge from cognitive psychology research, and more specifically from work on human contingency learning. Contingency learning refers to the basic human ability to learn the relationship between two or more 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 MacLeod, 2019; Schmidt, 2021). In an incidental learning procedure, the participant is not given the explicit goal to learn a regularity. Rather, the participant is asked to engage in one task (e.g., identify a target stimulus), but a regularity exists in the task (e.g., an informative secondary stimulus or a predictable sequence of stimuli) that, if learned, allows for anticipation of the likely response. We want to specify that here we used the term “incidental” because we refer to the acquisition of new information without the goal to learn (Kerka, 2000). We note that a separate (albeit correlated) issue from the incidental (vs. deliberate) nature of learning is whether participants are aware of what they have learned. For decades, there has been a heated debate about the nature (implicit or explicit) of the knowledge acquired through “implicit” or incidental learning (Cleeremans et al., 1998). Although we will take some measures of awareness in the present report, it is not our goal to discuss this debate in any detail.

Previous research suggests that learning the relationship between events occurs automatically, that is, people are sensitive to frequency of occurrence information (Zacks & Hasher, 2002) and to probabilistic patterns (Kelly & Martin, 1994), and simply attending to events is enough for activating learning of the co-occurrence of these events. Furthermore, people are not just sensitive to the co-occurrences around them, but they can learn this information and use it in a variety of tasks (e.g., in language acquisition; see Aslin et al., 1998; Saffran, Aslin, et al., 1996; Saffran et al., 1997; Saffran, Newport, et al., 1996). We note that we not only have a natural sensitivity in detecting the frequency and probability of events, but this sort of incidental learning can also occur very quickly. Indeed, many learning procedures, such as sequence learning (Nissen & Bullemer, 1987; Turk-Browne et al., 2005), 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), produce a rapid learning effect.

We took particular inspiration from the color-word contingency learning procedure of Schmidt et al. (2007; for related learning procedures, see Carlson & Flowers, 1996; Miller, 1987; Mordkoff & Halterman, 2008; Musen & Squire, 1993). Similar to the color-word Stroop procedure (Stroop, 1935), participants are asked to respond to the color of words by pressing a corresponding button, while ignoring the words. However, the words are neutral (unlike the Stroop) and to induce the acquisition of the contingencies, the words are presented most often in one color (e.g., “move” most often in blue) and rarely in the other colors (“move” rarely in red). Although participants are not informed of the contingencies between colors and words and often do not become aware of the manipulation, they respond quicker and more accurately to *high-contingency* trials, where the word is presented with the expected color (e.g., “move” in blue), than to *low-contingency* trials, where the word is presented with an unexpected color (e.g., “move” in red; Schmidt & De Houwer, 2012b). This contingency learning effect can be explained by the greater familiarization with frequently-presented high-contingency trials relative to the rarely-presented low-contingency trials (Schmidt & De Houwer, 2016a). The learned regularities allow participants to anticipate the responses based on the presented words (Schmidt et al., 2007), thereby facilitating performance if the anticipated high-contingency response is, in fact, required. Interestingly, this effect is extremely robust, with essentially all participants showing a numerical effect, and it is acquired almost instantaneously from the start of acquisition (Lin & MacLeod, 2018; Schmidt et al., 2010; Schmidt & De Houwer, 2016).

A major part of the reason *why* learning is so rapid in this type of incidental learning procedure is probably due to the fact that participants see a very large number of trials in which a stimulus is presented and they rapidly respond to it. In other words, such procedures allow participants to cram substantial amounts of practice with novel stimuli into a very short time period (e.g., several hundred trials in a 10-15 min). As previously indicated, this is one of the difficulties with training sight-reading: traditional practice does not involve seeing a large amount of novel materials in a short time period. In any case, given how rapid and easy it is to learn with this type of incidental learning procedure, a similar approach might be equally effective in the automatization of visuomotor integration for sight-reading performance. In particular, we hypothesize that participants may be able to acquire the associations between note positions and note names, along with the corresponding actions (i.e., which note to play) with similar efficiency. Indeed, learning in this type of incidental learning procedure primarily involves the learning of the association between the task-irrelevant stimulus (in the experiments to be described shortly: the note position) and the response to make (e.g., the key to press on a keyboard), or stimulus-response learning (Geukes et al., 2019; Miller, 1987; Schmidt et al., 2007; Schmidt & De Houwer, 2012a, 2016a). This is particularly interesting in the context of sight-reading, where automatization of the association between the note position and the action to perform on the instrument is needed. Our studies will therefore follow a similar logic as the color-word contingency learning described above, but with musical materials.

We note that incidental or implicit learning tasks have been used to investigate the learning of music materials in prior work. However, this prior work involved the learning of music that we listen to. For instance, many authors studied the implicit acquisition 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). In particular, the participants listen to music sequences and the learning of the structures underlying these sequences is then tested. However, the role of implicit or even incidental procedures in acquiring music skills useful for performance (e.g., how to play) is not clear yet.

The current research

Our adapted musical contingency-learning procedure is a hybridization of the above-mentioned musical Stroop and color-word contingency learning procedures. Our task follows the same structure of the musical Stroop task of Grégoire et al. (2013), in which a note is presented on a musical staff, which we will refer to as the note-position or simply the note. Written inside the note is the name of a note (e.g., "mi"), or note-name. Critically, as illustrated in Figure 4, the note-name can be either congruent with the position of the note (e.g., "ré" written inside the note for "ré") or incongruent (e.g., "mi" written inside of the note for "la"). However, to induce the learning of the note-name/note-position associations, our task follows the same logic as the color-word contingency learning procedure of Schmidt et al. (2007). In Experiment 1, each note was presented much more frequently with the congruent note-name (18 of 24 presentations, or 75%) than with any of the incongruent note-names (6 of 24 presentations, or 25%). For instance, the note-position for "do" was presented much more often with the note-name "do" than with the note-names "ré", "mi", and so on.

Participants simply respond as quickly and accurately as possible to the task-relevant stimulus (note-name) while ignoring the task-irrelevant stimulus (note-position). Critically, the note-position is informative in our adaptation (i.e., the note-position is predictive of the probable correct response to the note-name). Thus, learning could occur incidentally, and nonmusicians could learn the keyboard actions to perform for the note positions via the contingencies between the note-positions and responses to the note-names. We note that we use an imperfect contingency manipulation (i.e., not all trials are congruent) because this allows us to measure learning while it is occurring (i.e., by contrasting performance on high- and low-contingency trials; see Discussion for further remarks on this point).

Previously in the introduction, music sight-reading has been defined as a transcriptional task, where music symbols are translated into motoric actions (Sloboda, 1982, 1985). To study closely the acquisition of this task, we required our participants to respond to the note-names by pressing an assigned key on a computer keyboard. This type of arbitrary stimulus-response assignment is similar to the learning of playing a new musical instrument, where, for instance, a novice musician must learn which keys to press on a piano keyboard for each note.

It was anticipated that our incidental learning procedure would allow for rapid automatization of sight-reading skills, primarily because participants can experience a relatively large number of randomized trials with the congruent correspondences between note-positions and the keyboard responses to note-names. However, this is not to say that the deliberate intention to learn will not aid learning further. Schmidt and De Houwer (2012a, 2012d) compared the performance in the color-word contingency learning procedure between a deliberate learning group (which was informed of the contingencies present in the paradigm) and an incidental one (which was not informed of the contingencies). Their results showed better performance for the deliberate learning group, suggesting that intentionality plays a role in learning the contingencies (for a similar result in sequence learning, see Destrebecqz, 2004). Therefore, to assess the role of intentionality during learning, in Experiment 1, participants were divided into a deliberate learning group, instructed to pay attention to the contingencies, and an incidental learning group, who received no instructions about the presence of contingencies. It was hypothesized that, most critically, even the incidental learning group would show evidence of learning. However, the deliberate learning group might show even more robust learning.

In addition, subjective and objective awareness measures (see Cheesman & Merikle, 1984) were taken to assess the verbalizable knowledge of the contingencies acquired by participants. *Subjective awareness* is measured by simply asking participants whether they noticed the contingent regularities. *Objective awareness* is measured by asking participants to forced-choice guess the “name” of each note-position, with awareness indicated by above-chance guessing. The objective awareness test also serves as a “test” phase of verbalizable knowledge of the meaning of the note positions.

To summarize, we hypothesized that the incidental contingency learning procedure will help nonmusicians to easily learn the visuomotor translation of music symbols. However, based on previous research (Schmidt & De Houwer, 2012a, 2012d), it is expected that the deliberate intention to learn can help learning even further. Moreover, in a long-term perspective, this research aims to provide the starting point to create a tool that allows nonmusicians (or even experienced musicians) to learn (or improve) sight-reading abilities.

Pilot study

In the interest of full disclosure, we note that we initially conducted a pilot study with 41 participants (undergraduate psychology students from the University of Burgundy). The pilot was identical to Experiment 1 below, with two exceptions. First, there was no deliberate learning group (i.e., all participants learned incidentally). Second, the contingency manipulation was much weaker. Specifically, each note was presented only six times more frequently with the congruent note-name than any of the incongruent note-names (instead of 18 times more frequently in Experiment 1), meaning that congruent pairings occurred on only 50% of trials.

The resulting contingency effect was not significant in response times (RTs), $t(40) = 1.29$, $p = .205$, $d = -.201$, $BF_{10} = .364$, or errors, $t(40) = -1.32$, $p = .195$, $d = .206$, $BF_{10} = .377$, but the difference between low-contingency and high-contingency trials ($M_{\text{low-high}} = 8.28$, $SD = 41.1$) in RTs was encouraging. We thus strengthened the contingency manipulation in

Experiment 1, as this should increase the size of the learning effect. For instance, Forrin and MacLeod (2018) showed that the magnitude of the color-word contingency effect is exponentially related to contingency strength. That is, the effect gets much larger the stronger the contingency manipulation is.

Thus, for the present study we decided to (a) increase the strength of the contingency manipulation to elicit a larger congruency effect, (b) increase the sample size for more statistical power, and (c) introduce a deliberate learning group to explore the role of intentionality in a musical notation acquisition context. Supplementary material on our pilot experiment can be obtained by following the link: https://osf.io/fzex7/.

Experiment 1

In Experiment 1, two main hypotheses are investigated: 1) Based on color-word contingency learning research, it is expected that after a very small amount of practice, nonmusicians should incidentally learn which note-name corresponds to which note-position, and should therefore respond faster to the high-contingency (or "congruent") pairings relative to the low-contingency (or "incongruent") pairings, and 2) after a short learning phase, both the participants in the deliberate and incidental learning groups will be able to explicitly read musical notation, performing above chance in the objective awareness test phase.

Method

Participants

We recruited 123 undergraduate psychology students at the University of Burgundy. The participants received course credits for their voluntary participation. Participants were randomly assigned to the deliberate and incidental learning groups. Sixty-two participants (deliberate learning group) were asked to focus on the contingencies occurring during the learning phase. The remaining 61 participants (incidental learning group) did not receive any instructions about the contingencies present in the task. Our inclusion criteria were not being a musician and not being able to read musical notation. These inclusion criteria were indicated in the recruitment advertisement. All the procedures were conducted in accordance with the Declaration of Helsinki. A written consent form was signed by all participants before beginning the study. Participants' anonymization was guaranteed.

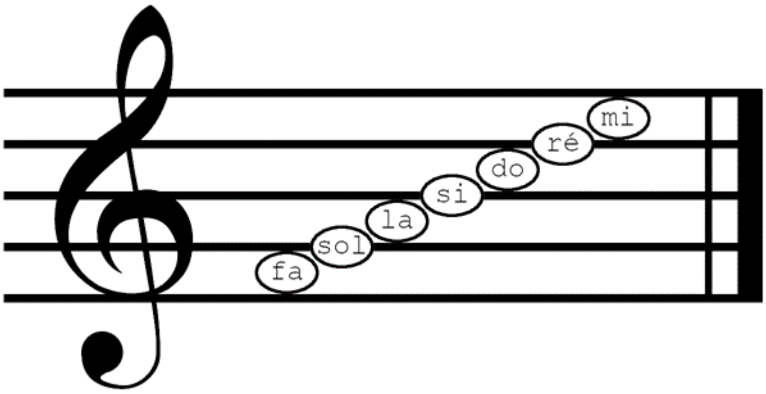
Apparatus, Design, and Procedure

In addition to the contingency manipulation, we made some additional changes to the typical musical Stroop procedure to aid learning. First, the musical staff was presented in the center of the screen in one fixed position. In the original experiments of Grégoire et al. (2013), the location of the staff was pseudorandomly varied in the four corners of the screen to prevent iconic memory of the staff. For the present report, however, we were actively aiming to train participants to learn location-to-response correspondences (i.e., note-position to note-name correspondences), so a fixed staff location was deemed desirable. Additionally, the note-position was presented slightly in advance of the note-name. This was done because it is known that advanced presentation of predictive cues boosts learning (Schmidt & De Houwer, 2016), likely because this gives the cue a "head start" to influence identification of the target. Finally, we used manual (key press) responses rather than oral naming responses. This was done, in part, for convenience and, in part, because a less automatic response modality (i.e., arbitrary stimulus-key assignments are slower than simple reading) allows more time for the cue (note-position) to influence responding to the target (Forrin & MacLeod, 2017; Schmidt, 2018). Moreover, as already suggested in the introduction, arbitrary stimulus-key assignments are similar to the motoric response that novice musicians practice when learning to play an instrument.

The experiment was programmed and ran with E-Prime 2 (Psychology Software Tools, Pittsburgh, PA) and run on laptops with 1080p resolution. During the main parts of the experiment, participants responded with the Z-I keys on a standard AZERTY keyboard. The keys 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). 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.

For stimuli, we used the seven notes from one octave (excluding the repeated octave) but beginning from "fa" (F4) and ending at "mi" (E5), as illustrated in Figure 5. We selected notes from "fa" to "mi" simply to use notes that fit within the main treble staff (i.e., first to fourth space). For instance, the first "do" (C4) falls on one line below the staff and normally is marked with a small strikethrough to indicate the position, which was deemed undesirable. In French, the note names are "do," "ré," "mi," "fa," "sol," "la," and "si", unlike in English where A-G letter names are typically used. All target stimuli were presented in black 30 pt. Courier New font on a white screen, unless otherwise noted.

Figure 5 - Full range of note positions used in the experiments, with congruent names.



Note. An individual note was horizontally centered on a smaller staff in the actual experiment, as in Figure 4.

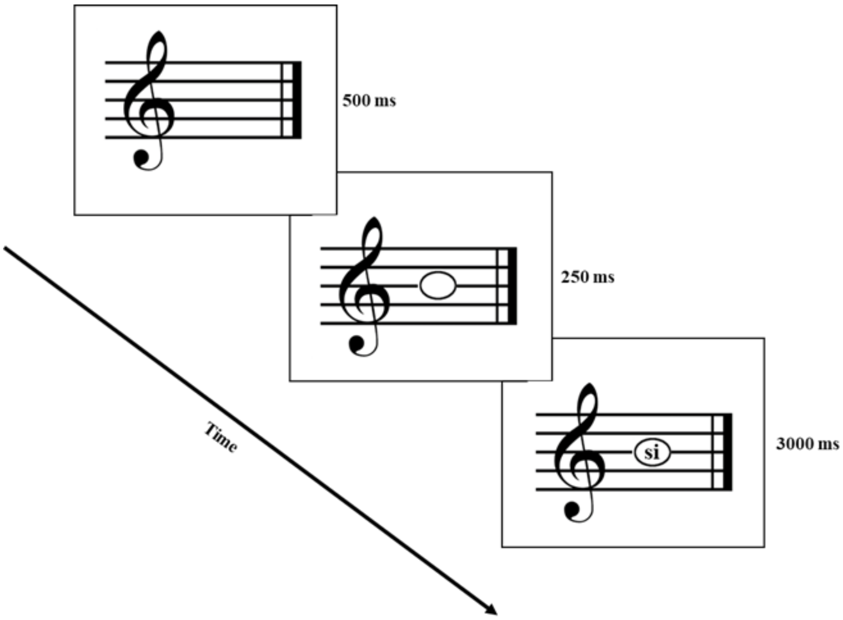
The experiment involved five phases. The goal of the first two phases was to allow participants to practice and automatize the note name-to-key assignments before proceeding to the actual learning phase. Results for these phases are not analyzed. In these practice phases, participants were not presented with notes or the musical staff, but only the written note names. In the first of these phases, the trial started with a fixation cross ("+") in the center of the screen for 500 ms. This was followed by one of the seven the French note-names (*fa*, *sol*, *la*, *si*, *do*, *ré*, or *mi*) presented in the center of the screen until the participant pressed the corresponding response key (no time limit). Following correct responses, the next trial began immediately. Following incorrect responses, the note-name changed color to red (255,0,0; or E-Prime/HTML "red") and stayed on the screen until the participant pressed the correct key. During the entire trial, the seven key labels (*fa* through *mi*), corresponding to the keyboard response keys, were presented at the bottom of the screen in bold 18 pt. Courier New font with five spaces between each, *x*-axis centered and below the target (centered at 600 px. on the *y*-axis). No specific instructions were given on how to use the keyboard responses. Each of the seven note names was presented once per block in random order, with ten blocks total (70 trials). The second practice phase was identical in all aspects, except that the on-screen key reminder was removed, and participants were encouraged to try to respond from memory (though the keys on the keyboard remained labelled in case the participant was particularly lost).

After these two training phases, and to study whether differences occurred between deliberate and incidental leaning, we added an extra instruction screen before the learning phase for half of the participants (deliberate learning group), which instructed them about the contingency manipulation and asked them to try to learn the contingencies. The instructions were (translated from the French version):

Note: Each note will be presented more frequently with the correct note name and less frequently with the incorrect note names. Try to learn the note name for each note position.

The following third phase was the main learning task used to assess learning in response times and errors. On each trial, participants were presented with the musical staff (see Figure 4), an image of 602 × 909 px. (squished slightly to 602 × 902 px. to better align notes with the staff), which remained centered on the screen throughout the whole trial. At the start of the trial, the empty staff was presented for 500 ms. The note (67 × 100 px.) was then added to the staff for 250 ms, *x*-axis centered at 800 px. and *y*-axis centered either on or between one of the lines for the given note-position (522, 482, 442, 402, 362, 322, or 282 px.). The note-name was then written inside the note and participants had 3000 ms to respond. The entire procedure for stimuli appearance during the learning phase is illustrated in Figure 6.

Figure 6 - Order of stimuli appearance during the learning phase.



If the participant responded correctly, the next trial began immediately. If they responded incorrectly or failed to respond in 3000 ms, the note name was replaced with "XXX" in red for 500 ms. During the learning phase, there were two blocks of 168 trials (336 trials in total), each randomly ordered (without replacement) and there was no break between the two blocks. Each note was presented 18 times per block with the congruent note-name (e.g., “fa” in the note for fa) and once each with the remaining six note names (e.g., “fa” in the note for do). Thus, each congruent name-note pairing was more frequent (high-contingency) than each incongruent name-note pairing (low-contingency). The congruency (or contingency learning) effect was measured as the difference between low- and high-contingency RTs (and errors).

Following the main learning phase, we additionally collected contingency awareness data to assess the awareness of participants of the contingency manipulation in the final two phases. In particular, participants were assessed for both subjective and objective awareness (Cheesman & Merikle, 1984). Respectively, *subjective awareness* is defined as a participant’s ability to verbally describe their experience, while *objective awareness* is defined by a participant’s ability to discriminate (e.g., better-than-chance guessing) between experienced and unexperienced events. For the *subjective awareness* measure (i.e., the fourth phase), the on-screen instructions told participants (translated from French):

> *During the third part of this experiment, note names were written inside the notes. Each note was presented more frequently with one note name than the others. That is to say, one note was frequently presented with “do,” another frequently with “re,” etc. Did you notice these regularities?*

Participants could respond “yes” or “no” with a key press.

Directly after, we introduced the *objective awareness* measure test phase (i.e., the fifth and final phase) as a more explicit test of verbalizable knowledge to (a) test whether the association between note-position and note-name was acquired, and (b) investigate whether the information incidentally acquired could be expressed explicitly. The phase began with the following instructions (translated from French):

Now, the task is similar, except that you will only see a note (not a note-name). Try to guess the name of the note by pressing the appropriate key on the keyboard.

The task was similar to the learning phase, except that (a) only the note-positions (without note-names) were presented, (b) the on-screen key reminder was re-added below the musical staff (y-axis centered on 775 px.), (c) there was no time limit to respond, and (d) there was no accuracy feedback. Thus, participants had to respond to the notes themselves (previously task irrelevant) rather than to the note-names. There were three blocks of each of the seven notes (21 trials in total), presented randomly without replacement.

Data Analysis

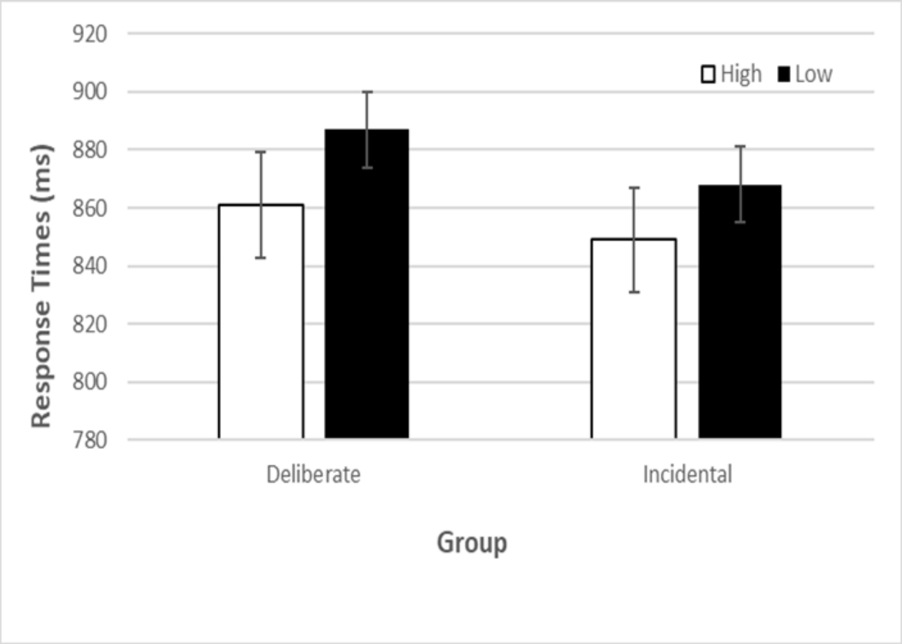
Analyses of the learning phase were conducted on mean correct RTs and error rates. Trials in which participants failed to respond in 3000 ms (i.e., the response deadline) were eliminated. Repeated measures ANOVAs for RTs and for error rates were conducted to assess the overall main effects of contingency, instruction, and the interaction between them. Furthermore, we ran another repeated measures ANOVA for RTs and error rates with the added factor of block (Block 1 and Block 2) to assess the presence of a contingency effect from the start of the acquisition process. If this is the case, we expect no significant interaction between blocks and contingency. One-sample t tests were used to assess learning rates between the groups. Pearson’s correlations were performed to assess relations between objective and subjective awareness and the contingency effect. All analyses were evaluated at the $\alpha = .05$ level of significance. Additionally, we estimated the Bayes factor for all the data using JASP software (JASP Team, 2019). All the Bayesian analyses were done using the standard noninformative Cauchy prior in JASP with a default width of 0.707. A BF_{10} between 3 and 10 allows us to conclude that we have moderately strong evidence for H_1. The data set and R script are available via the following link: https://osf.io/fzex7/.

Results

Response Times

The RT results for Experiment 1 are presented in Figure 7. A repeated measures ANOVA for RTs with the factors Contingency (high vs. low) and Group (deliberate vs. incidental) showed a significant main effect of Contingency, $F(1,121) = 32.347$, $p < .001$, $\eta^2 = .211$, $BF_{10} > 100$, indicating faster responses for high-contingency trials ($M = 855$ ms, $SD = 112$) than for low-contingency trials ($M = 877$ ms, $SD = 115$).

Figure 7 - Experiment 1, RTs for deliberate and incidental groups.



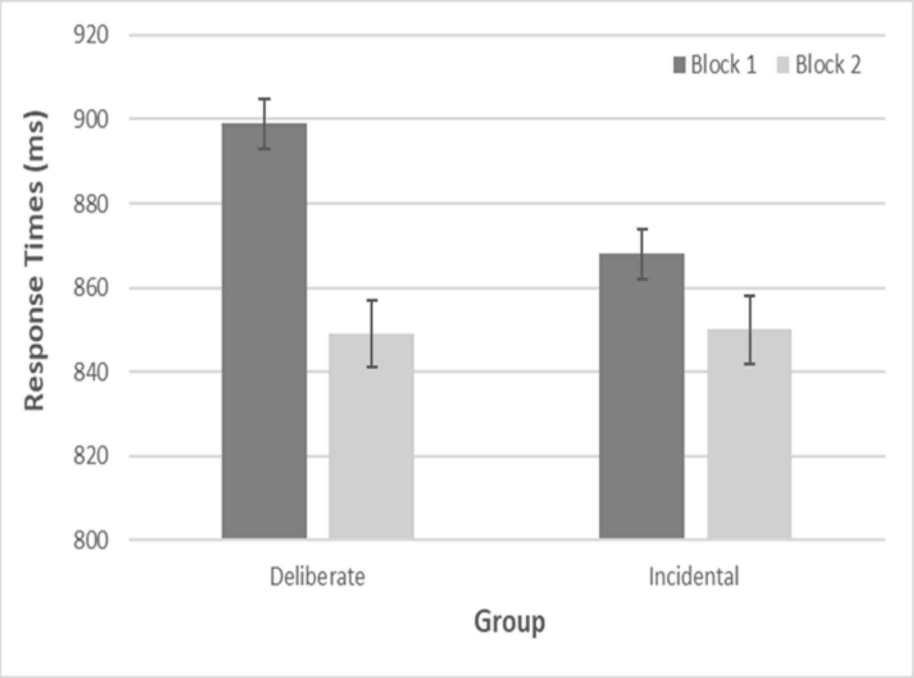
Note. Interaction between Contingency (High and Low) and Group (Deliberate and Incidental), standard error bars are shown in the figure.

The main effect of Group was not significant, $F(1,121) = .580$, $p = .448$, $\eta^2 = .005$, $BF_{10} = .554$. Interestingly, the interaction between Contingency and Group was also not significant, $F(1,121) = .797$, $p = .374$, $\eta^2 = .007$, $BF_{10} = .278$, indicating no significant differences between deliberate vs. incidental learning groups for the contingency effect, although the effect was numerically larger for the deliberate learning group (see Figure 7; $M_{\text{low-}}$

$_{\text{high}}$ = 25.7, *SD* = 38.4; *t*(61) = 5.25, *p* < .001, *d* = .667, BF_{10} > 100) compared to the incidental one ($M_{\text{low-high}}$ = 18.7, *SD* = 47.6; *t*(60) = 3.07, *p* =.003, *d* = .393, BF_{10} = 9.320).

Additionally, a repeated measures ANOVA for RTs with the factors Block (1 vs. 2), Contingency (high vs. low), and Group (deliberate vs. incidental) was computed to analyze the data for rapid acquisition of the contingencies and possible differences across blocks between the two groups. A significant main effect for Blocks was found, *F*(1,121) = 44.053, *p* < .001, η^2 = .267, BF_{10} > 100, showing significantly faster RTs in Block 2 (*M* = 849 ms, *SD* = 117) compared to Block 1 (*M* = 884 ms, *SD* = 115), indicating a standard practice effect on mean RT. The main effect for Contingency was also significant, *F*(1,119) = 32.363, *p* < .001, η^2 = .211, BF_{10} > 100.

Figure 8 - Experiment 1, interaction between Block and Group.



Note. Averaged response times across high and low contingency trials for block (Block 1 and Block 2) for the deliberate and incidental learning groups (standard error bars are shown).

Block and Contingency did not interact, $F(1,121) = .543$, $p = .463$, $\eta^2 = .004$, $BF_{10} = .277$, suggesting that the learning of contingencies is fast rather than appearing gradually across blocks. On the other hand, the interaction between Block and Group was significant, $F(1,121) = 9.839$, $p = .002$, $\eta^2 = .075$, $BF_{10} = 95.284$. Specifically, as illustrated in Figure 8, we found a significant difference in RTs for the deliberate learning group ($M_{\text{Block 1-Block 2}} = 50.65$; $t(121) = 6.939$, $p < .001$). This difference was not significant for the incidental learning group ($M_{\text{Block 1-Block 2}} = 18.14$; $t(121) = 2.465$, $p = .091$). Finally, the Contingency × Block × Group interaction was not significant, $F(1,121) = .279$, $p = .599$, $\eta^2 = .002$, $BF_{10} = .190$

Error Rates

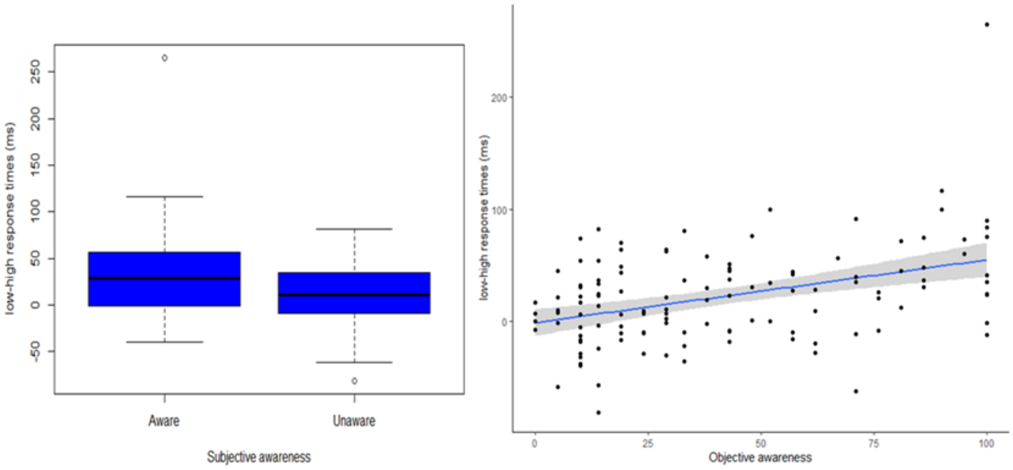
A repeated measures ANOVA for errors with the factors Contingency (high vs. low) and Group (deliberate vs. incidental) did not reveal a main effect of Contingency, $F(1,121) = .081$, $p = .776$, $\eta^2 = .001$, $BF_{10} = .145$, or Group, $F(1,121) = .115$, $p = .735$, $\eta^2 = .001$, $BF_{10} = .291$. The interaction between Contingency and Group was also not significant, $F(1,121) = .015$, $p = .901$, $\eta^2 = 0.00$, $BF_{10} = .186$ (deliberate learning group, $M_{\text{high}} = .976$ ms, $SD = .025$, $M_{\text{low}} = .990$ ms, $SD = .029$; incidental learning group, $M_{\text{high}} = .977$ ms, $SD = 0.027$, $M_{\text{low}} = .980$ ms, $SD = .025$). Given the lack of a contingency effect in errors, a block analysis was not performed.

Subjective and objective awareness

For the subjective awareness question, 33 of 62 participants (53%) in the deliberate learning group reported that they noted the regularities, and 27 of 61 participants (44%) in the incidental learning group. Subjective awareness rates were not significantly different between the two groups, $M_{\text{deliberate-incidental}} = 9\ \%$, $t(121) = .990$, $p = .324$, $d = .179$, $BF_{10} = .300$ (deliberate learning group: $M = 53\%$; incidental learning group: $M = 44\%$).

Using one-sample *t* tests, we found that the rates of objective awareness (test phase accuracy) were above chance (1/7 or 14.3%) in both groups: deliberate learning group (M = 50.6%, SD = 31.1), $t(61) = 9.19$, $p < .001$, d = 1.17, $BF_{10} > 100$, incidental learning group (M = 32.0%, SD = 27.7), $t(60) = 4.98$, $p < .001$, d = .637, $BF_{10} > 100$. Objective awareness was higher for the deliberate learning group than for the incidental learning group, and a Welch two-sample *t* test showed that this 19% differences between the two groups was significant, $t(120) = 3.51$, $p = .001$, d = .633, $BF_{10} = 42.530$.

Figure 9 - Experiment 1, correlations between contingency effect and subjective and objective awareness.



Note. In the left panel, the correlation between the contingency effect and subjective awareness is shown. In the right panel, the correlation between the contingency effect and objective awareness (test phase) is shown.

The RT-contingency effect (i.e., low- minus high-contingency) correlated significantly with both subjective awareness, $r(121) = .239$, $p = .008$, $BF_{10} = 3.760$, and objective awareness, $r(121) = .401$, $p < .001$, $BF_{10} > 100$, as shown in Figure 9. Additionally, the contingency effect was significant for both participants who were subjectively aware (M = 32.7, SD = 48.1), $t(59) = 5.28$, $p < .001$, d = *.681*, $BF_{10} > 100$, and for those who declared to be unaware (M = 12.2, SD = 35.5), $t(62) = 2.72$, $p = .009$, d = *.342*, $BF_{10} = 3.941$, suggesting stronger contingency effect for participants with greater awareness. For the objective awareness factor, we also computed the regression intercept at chance guessing (Greenwald et al., 1995). That is, we calculated a regression with objective awareness as the predictor and the RT contingency effect as the dependent variable. Objective awareness was re-centered at chance guessing (1/7, or 14.3%). The intercept therefore indicates the size of the contingency effect when participants are guessing at chance in the objective awareness phase. This intercept was numerically above zero in the sample as a whole (intercept M = 6.989), but not significantly, $t(121) = 1.46$, $SE = 4.78$, $p = .146$, $BF_{10} = 1.0$. Globally, the data show an impact of contingency knowledge on the size of the RT contingency effect, though it remains unclear whether and to what degree implicit learning also contributes to the effect. In contrast to the RT data, the error contingency effect (low- minus high-contingency errors) was not correlated with subjective awareness, $r(121) = -.018$, $p = .845$, $BF_{10} = .115$, or objective awareness, $r(121) = .001$, $p = .993$, $BF_{10} = .113$, which is not surprising given the lack of a significant contingency effect in errors.

Discussion

As hypothesized, in Experiment 1 we found a contingency effect, suggesting that nonmusicians were able to incidentally learn the associations between note-positions and the keyboard responses to note-names. Furthermore, in line with previous research, the block analysis suggests a rapid acquisition of the contingencies starting from the beginning of the learning phase. Although both groups responded significantly higher than chance in the objective awareness phase, the deliberate learning group was more accurate than the incidental one. This result may indicate an influence of attention in explicitly reporting the new acquired information. Overall, a relationship between the contingency effect and awareness was revealed by the significant correlations.

Experiment 2

Experiment 2 addresses a potential caveat with Experiment 1. It may be argued that the contingency effect in Experiment 1 can be due to previous implicit knowledge about note-name/note-position associations, rather than contingency learning. Although participants claimed that they were not able to read music notation, it is possible that they studied music at school and remember more than they imagined or even that some musicians misrepresented their music reading abilities in order to participate. If this were true, then it could be the case that no actual learning occurred in Experiment 1. Therefore, to address this concern and to also investigate whether previous musical knowledge influences the effect, we ran a second experiment. Experiment 2 was identical to Experiment 1, except that the high-contingency pairings were no longer the congruent pairings. Each note position was presented more often with one of the six incongruent note names (e.g., “ré” written inside the note for “fa”) on high-contingency trials and rarely with the remaining congruent and incongruent names (e.g., “ré” written inside the note for “ré”) on low-contingency trials. Participants were divided in six groups, such that across participants every note position *except* the congruent note was high contingency for a given note name.

Unlike Experiment 1, in Experiment 2 contingency was different from congruency. The congruent trials (e.g., “ré” written inside the note for “ré”) were presented much less often than the high-contingency incongruent trials (e.g., “ré” written inside the note for “fa”). Therefore, in Experiment 2 we speak about both the contingency effect (low- minus high-contingency trials) and the congruency effect (incongruent minus congruent trials). If previous musical knowledge is not present, the scrambling of the note-name to note-position associations should not be important, and we anticipate that participants will show a contingency effect similar to the one found in Experiment 1 (faster RTs for high-contingency trials compared to low-contingency trials). On the contrary, if participants possess undisclosed previously acquired musical knowledge, then we should anticipate a congruency effect (faster RTs for congruent trials compared to incongruent ones) despite the high- vs. low-contingency presentation. Of course, it is also possible that both effects will be observed: a true learning effect within the experiment in addition to a congruency effect due to undisclosed sight-reading knowledge.

Method

Participants

Experiment 2 took place during the Covid-19 pandemic, so to adhere to the general health recommendations that restricted the possibility to recruit new participants to come to the lab, we ran Experiment 2 using the online Prolific.co platform. 132 participants clicked to start the experiment, but we excluded participants who abandoned the experiment before completion or did not actually begin the task. The remaining 60 participants, who received monetary compensation (£2) for their participation, were randomly assigned to each of the six scrambled note-name/note-position conditions, described below. Each condition was composed of 10 participants. The inclusion criteria were the same as those used for Experiment 1 and they were mentioned in the recruitment advertisement. All the procedures were conducted in accordance with the Declaration of Helsinki. A consent form was signed by all participants before beginning the study. Participants' anonymization was guaranteed.

Apparatus, Design, and Procedure

The experiment was programmed and run with Psytoolkit (Stoet, 2010, 2016). The structure of Experiment 2 was basically the same as Experiment 1, with the following exceptions. All participants learned incidentally, as in the incidental learning group of Experiment 1. Thus, no instruction about the contingencies was given. In the learning phase, we used scrambled note-name/note-position associations. That is, each note was presented 18 times per block with one of the incongruent note-name pairing (e.g., "ré" in the note for "fa") and once each with the remaining congruent and incongruent note-name pairings (e.g., "ré" in the note for “ré” and “ré” in the note for “do”). Thus, one specific incongruent name-note pairing was more frequent (high-contingency) than the congruent and each remaining incongruent name-note pairings (low-contingency). We created six groups by shifting the name-position correspondences by 1, 2, 3, 4, 5, or 6 tones (e.g., the position “ré” most often with the name “mi”, “mi” most often with “fa”, etc. for Group 1; the position “ré” most often with “fa”, “mi” most often with “sol”, etc. in Group 2; etc.). Thus, across participants all note positions were high-contingency with each note name, except the congruent pairing.

Data Analysis

The same data analysis criteria used in Experiment 1 were applied for Experiment 2 with some exceptions: no block ANOVA was assessed; no instruction factor was used since all participants learned incidentally in Experiment 2. However, we ran a repeated measures ANOVA with Congruency as factor to evaluate the influence of congruent vs. incongruent trials on the learning process. High-contingency incongruent trials were eliminated from this analysis, so that the low-contingency congruent trials were compared only to low-contingency incongruent trials. The data set and R script are available via the following link: https://osf.io/fzex7/.

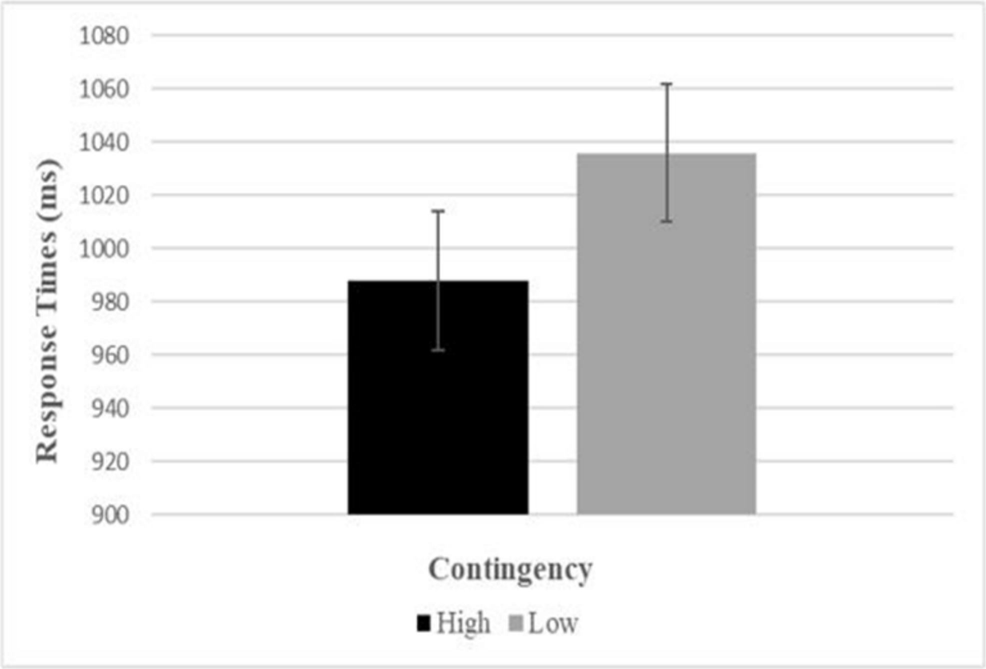
Results

Response Times

The RT results for Experiment 2 are presented in Figure 10. The repeated measures ANOVA for RTs with Contingency (high vs. low) and Group (1, 2, 3, 4, 5, 6) as factors, showed a significant main effect of Contingency, $F(1,54) = 55.284$, $p < .001$, $\eta^2 = .506$, $BF_{10} > 100$ (high-contingency trials, $M = 988$ ms, $SD = 207$; low-contingency trials, $M = 1036$ ms, $SD = 206$). The main effect of Group was not significant, $F(5,54) = 1.05$, $p = .400$, $\eta^2 = .088$, $BF_{10} = .527$, and the interaction between Contingency and Group was also not significant,

$F(5,54) = .565$, $p = .726$, $\eta^2 = .050$, $BF_{10} = .064$, suggesting no differences between groups for the contingency effect.

Figure 10 - Experiment 2, averaged mean for contingency effect.



Note. Averaged mean scores between groups for high- and low-contingency trials. Error bars represent standard errors.

Interestingly, when using Congruency (congruent vs. incongruent) and Group (1, 2, 3, 4, 5, 6) as factors, the repeated measures ANOVA for RTs showed a significant main effect of Congruency, $F(1,54) = 4.668$, $p = .035$, $\eta^2 = .080$, $BF_{10} = 1.598$ (congruent trials, $M = 996$ ms, $SD = 201$; incongruent trials, $M = 1045$ ms, $SD = 219$). The main effect of Group was not significant, $F(5,54) = 1.55$, $p = .190$, $\eta^2 = .126$, $BF_{10} = .674$, as was the interaction between Congruency and Group, $F(5,54) = .411$, $p = .839$, $\eta^2 = .037$, $BF_{10} =.100$.

Error Rates

The repeated measures ANOVA for errors with Contingency (high vs. low) and Group (1, 2, 3, 4, 5, 6) as factors did not reveal a main effect of Contingency, $F(1,54) = 1.267$, $p =$

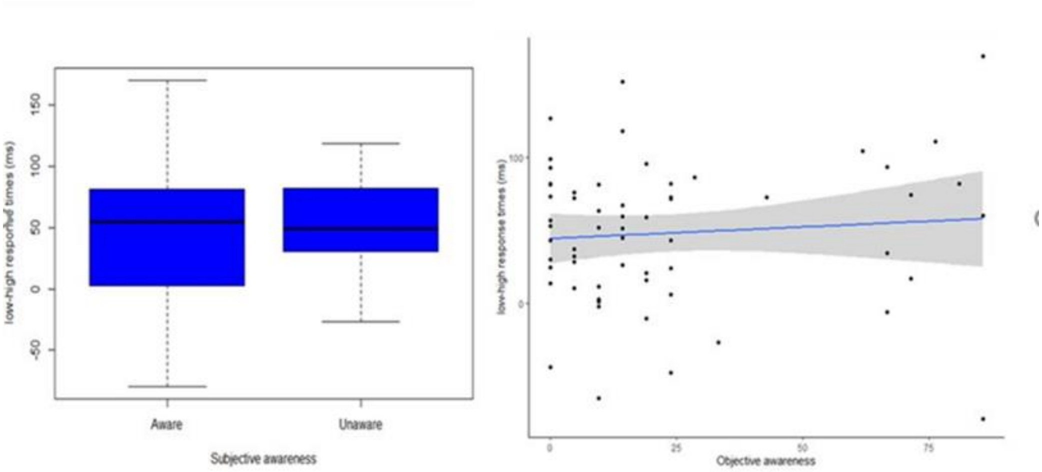
.265, η^2 = .023, BF_{10} = .329 (M_{high} = 9.74%, *SD* = 8.70; M_{low} = 10.3%, *SD* = 10.5), or Group, *F*(5,54) = 1.17, *p* = .335, η^2 = .098, BF_{10} = .442. The Contingency by Group interaction was also not significant, *F*(5,54) = .875, *p* = .504, η^2 =.075, BF_{10} = .137.

Surprisingly the repeated measures ANOVA for errors with Congruency (congruent vs. incongruent) and Group (1, 2, 3, 4, 5, 6) as factors showed a significant main effect of Congruency, *F*(1,54) = 6.54, *p* = .013, η^2 = .108, BF_{10} = 1.614 ($M_{\text{congruent}}$ = 7.86%, *SD* = 9.18; $M_{\text{incongruent}}$ = 10.8%, *SD* = 11.4). The main effect of Group was not significant, *F*(5,54) = .797, *p* = .557, η^2 = .069, BF_{10} = .095, nor was the Congruency by Group interaction, *F*(5,54) = 2.12, *p* = .078, η^2 =.164, BF_{10} = .674.

Subjective and objective awareness

In Experiment 2, more than 50% of the participants (34 of 60) reported to be aware of the regularities. Overall, the rates of objective awareness (test phase accuracy) were above chance (1/7 or 14.3%), (*M* = 23.2%, *SD* = 26.5) *t*(59) = 2.60, *p* = .006, *d* = .335, BF_{10} = 3.018. The correlations between the RT-contingency effect (i.e., low- minus high-contingency) and subjective awareness, *r*(58) = .123, *p* = .350, BF_{10} = .247, and objective awareness, *r*(58) = .085, *p* = .519, BF_{10} = .197, were not significant. Additionally, the contingency effect was significant for both participants who were subjectively aware (*M* = 42.9%, *SD* = 57.5), *t*(33) = 4.35, *p* < .001, *d* = .746, BF_{10} > 100, and for those who declared to be unaware (*M* = 55.0%, *SD* = 35.5), *t*(25) = 7.90, *p* < .001, *d = 1.55*, BF_{10} > 100.

Figure 11 - Experiment 2, correlations between contingency effect and subjective and objective awareness.



Note. In the left panel, the correlation between the contingency effect and subjective awareness is shown. In the right panel, the correlation between the contingency effect and objective awareness (test phase) is shown.

The congruency effect correlated significantly with subjective awareness, $r(58) = .345$, $p = .007$, $BF_{10} = 5.671$, but not with objective awareness, $r(58) = -.057$, $p = .668$, $BF_{10} = .176$, as shown in Figure 11. Moreover, the congruency effect was not significant for participants who were subjectively aware ($M = -2.14$, $SD = 158$), $t(33) = -.079$, $p = .938$, $d = -.013$, $BF_{10} =.148$, but was significant for those who declared to be unaware ($M = 116$, $SD = 165$), $t(25) = 3.56$, $p = .002$, $d = .699$, $BF_{10} = 23.940$. As for Experiment 1, we computed a regression intercept at chance guessing (Greenwald et al., 1995) with the objective awareness factor. The result showed a significant intercept above zero, $t(54) = 6.934$, $SE = 6.735$, $p <.001$, $BF_{10} = 1.0$; intercept $M = 46.707$, suggesting that implicit learning contributed to the contingency effect.

Not surprisingly, the error contingency effect (low- minus high-contingency errors) was not significantly correlated with subjective awareness, $r(58) = -.155$, $p = .238$, $BF_{10} = .318$, or objective awareness, $r(58) = .122$, $p = .353$, $BF_{10} = .245$. The error congruency effect was also not significantly correlated with subjective awareness, $r(58) = .217$, $p = .096$, $BF_{10} = .621$, or objective awareness, $r(58) = -.000$, $p = .998$, $BF_{10} = .161$.

Discussion

In Experiment 2, we again observed a contingency learning effect. Unlike in Experiment 1, however, the high-contingency pairings were (specific) incongruent pairings in Experiment 2. Thus, preexisting sight-reading knowledge could not have produced this contingency learning effect. Indeed, any preexisting knowledge would actually work *against* a contingency learning effect, as the congruent pairings were low contingency. Interestingly, we did also observe a congruency effect when comparing the congruent and incongruent low-contingency pairings. This is a bit surprising given that past reports have failed to observe a congruency effect in nonmusicians (e.g., Crump et al., 2012; Drost et al., 2005; Grégoire et al., 2013; Stewart, 2005). Similarly, we did not find a robust contingency effect for congruent pairings in our other studies with the present paradigm when the contingency manipulation was too weak (including our pilot study and data from one of the conditions of some of our follow-up work to the present report). The reason for this congruency effect is unclear. One possible interpretation is that some of the participants did have prior sight-reading knowledge and failed to disclose this, but Experiment 3 will explore this and another potential interpretation.

Subjective but not objective contingency awareness was poorly correlated with the contingency effect. Some evidence for implicit learning contributions to the contingency effect were observed, including a significant contingency effect for subjectively unaware participants and a significantly positive intercept in the objective awareness data, which contrasts slightly with the results of Experiment 1. Participants also guessed at above-chance rates the interpretations of the note positions. Of course, these were technically the incorrect note interpretations (i.e., consistent with the incongruent contingencies).

Experiment 3

As previously mentioned, we were surprised to find a significant congruency effect in Experiment 2. As mentioned above, this congruency effect may have been due to the inclusion of some participants that did have preexisting sight-reading knowledge that they failed to disclose (e.g., due to underestimation of their knowledge). However, there may be another explanation for the congruency effect that does not assume that some of the participants had preexisting knowledge. Indeed, it is possible that a congruency effect might be observed even if participants do not know the association between note names and note positions. Instead, there may have been an inherent spatial compatibility between the down-to-up organization of the note positions and the left-to-right organization of the response keys.

Previous research (Rusconi et al., 2006) showed the presence of a SMARC (Spatial-Musical Association of Response Codes) effect, defined by the authors as "a variant of the well-known orthogonal stimulus-response compatibility effect, that is a preferential mapping of spatially lower stimuli on left responses and higher stimuli on right responses" (Rusconi et al., 2006, p. 14). For the authors, the SMARC effect reflects the spatial coding of pitches, with the highest pitches represented on the right and the lowest pitches on the left.

Recently, Ariga and Saito (2019) showed the presence of a SMARC effect in the absence of pitch. Although, in their study there was no auditory stimulation, the effect was elicited by written pitch names alone for both trained musicians and musically naïve participants. Overall, this evidence suggested that the human cognitive system automatically codes pitches spatially.

Therefore, regarding our results, it is possible that the congruency effect could be explained by a natural inclination to spatially code pitches. Indeed, the spatially lowest note position (fa) corresponded to the leftmost response (Z) in our prior experiments moving up to the highest note position (mi) with the rightmost response (I). As such, it could be that participants responded faster to the congruent pairings not because they knew the interpretation of the note positions, but because of the spatial compatibility between the stimulus and response locations. To test this hypothesis, we ran a third experiment. It is worth noting that the aim of this third experiment is not to further investigate the contingency learning effect that we observed in the prior two experiments; rather, we aim to test whether the congruency effect found in Experiment 2 was due to preexisting sight-reading knowledge or to a SMARC-like compatibility effect.

Experiment 3 was identical to the previous two experiments, except that no contingency manipulation was used. Each note-name/note-position pairing was presented the same number of times. However, to test the hypothesis of the presence of the SMARC effect we distinguished between congruent trials, compatible trials, and control trials (see Table 1 and the method section for more details). In particular, the response options were reordered such that the congruent response was *not* spatially compatible with the note position. For instance, the bottommost stimulus location (fa) was *not* the leftmost response. Congruent trials were therefore the trials in which the note position was presented with the true note name (e.g., the position for fa presented with “fa”), compatible trials were not congruent but were spatially compatible (e.g., the position for fa with the note name “do”), and all remaining pairings were controls. If participants do not possess undisclosed previously acquired musical knowledge, then we should not find a congruency effect (faster RTs for congruent trials compared to control ones). If participants are influenced by spatial compatibility, however, then we might find a compatibility effect (RTs faster for the compatible trials compared to control trials).

Table 1 - Experiment 3, Musical Stroop contingency learning manipulation.

Note Name	**Note position**						
	Do	Ré	Mi	Fa	Sol	La	Si
Do	**3**	3	3	*3*	3	3	3
Ré	3	**3**	3	3	*3*	3	3
Mi	3	3	**3**	3	3	*3*	3
Fa	3	3	3	**3**	3	3	*3*
Sol	*3*	3	3	3	**3**	3	3
La	3	*3*	3	3	3	**3**	3
Si	3	3	*3*	3	3	3	**3**

Note. Numbers of repetition for each trial. Congruent trials in bold, compatible trials in underlining italic and control trials in standard font.

Method

Participants

Experiment 3 was coded using Psytoolkit (Stoet, 2010, 2016) and run using the online Prolific.co platform. One hundred and seventy-five participants clicked through to the link to the experiment on Prolific, but we again excluded participants that did not complete the study or actually begin it. 119 participants, who received monetary compensation (£2), took part in the experiment. The inclusion criteria were the same used for Experiment 1 and they were mentioned in the recruitment advertisement. All the procedures were conducted in accordance with the Declaration of Helsinki. A consent form was signed by all participants before beginning the study. Participants' anonymization was guaranteed.

Apparatus, Design, and Procedure

Experiment 3 was identical to Experiment 2 with the following exceptions. During the "learning phase", no contingency manipulation was used. That is, each note position was presented equally often with all of the note names. Thus, there was actually no regularity to learn in the present experiment. Instead, we manipulated spatial compatibility and congruency. To dissociate the two, we changed the order of the key mappings. While the down-to-up note positions still went from "fa" to "mi", the key mappings went from "do" to "si". In this way, the leftmost response (e.g., do) did not correspond to the bottommost note position (fa). None of the note positions were spatially compatible with the congruent response. Therefore, we distinguished between: (a) *congruent trials*, in which the note name (e.g., “do”) was written in the congruent note position (e.g., “do”), (b) *compatible trials*, where the note name was spatially compatible with the order of the key mapping (e.g., the leftmost note name “do” written in the bottommost note position “fa”), and (c) *control trials*, which were neither congruent nor spatially compatible (e.g., the leftmost note name “do” written in the topmost position “si”). Concretely, the responses were shifted three places to the left, but otherwise maintained the same relative order (i.e., do, ré, mi, fa, sol, la, si). Hypothetically, it would have been possible to create six such orders (e.g., analogous to Experiment 2). However, we opted for this single ordering because for many of the possible orders the congruent and spatial compatible responses would be very close to each other spatially. The particular response ordering that we used maximally separates the congruent and compatible responses. Furthermore, no subjective awareness phase was shown because of the lack of contingency manipulation. A phase effectively identical to the objective awareness phases of the previous experiments was still included but was no longer a true “awareness” phase (as there was no contingency to be aware of this time). We will therefore refer to this simply as the “test” phase.

Data Analysis

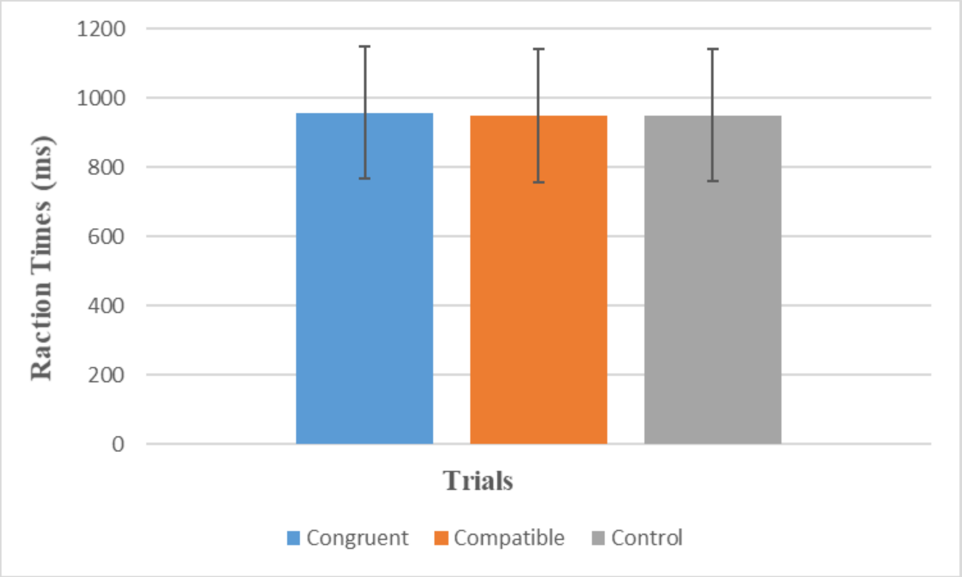
The same data analysis criteria as those used in Experiments 1 and 2 were applied in Experiment 3. We use *t* tests to compare RTs and error rates between the different trials: congruency, compatibility, and control. We also ran analyses on both the accuracy for congruency and the accuracy for compatibility in the test phase to study whether participants indicated the congruent and/or compatible responses more often than one would expect by chance. Given the absence of a contingency, participants should only indicate the congruent response more often than chance if they have preexisting sight-reading knowledge and should only indicate the compatible response more often than chance if they are influenced by spatial compatibility. The data set and R script are available via the following link: https://osf.io/fzex7/.

Results

Response Times

The *t*-tests analyses revealed no significant difference in RTs (Figure 12) between congruent and control trials ($M_{congruent\text{-}control}$ = 7.114, *SD* = 58.5), *t*(117) = 1.322, *p* = .189, *d* = .121, BF_{10} = .329, or between compatible and control trials ($M_{compatible\text{-}control}$ = -.755, *SD* = 58.9), *t*(117) = -.143, *p* = .887, *d* = -.013, BF_{10} = .104.

Figure 12 - Experiment 3, RTs for different trials.



Note. Mean RTs scores for the different trials: congruent, compatible and control.

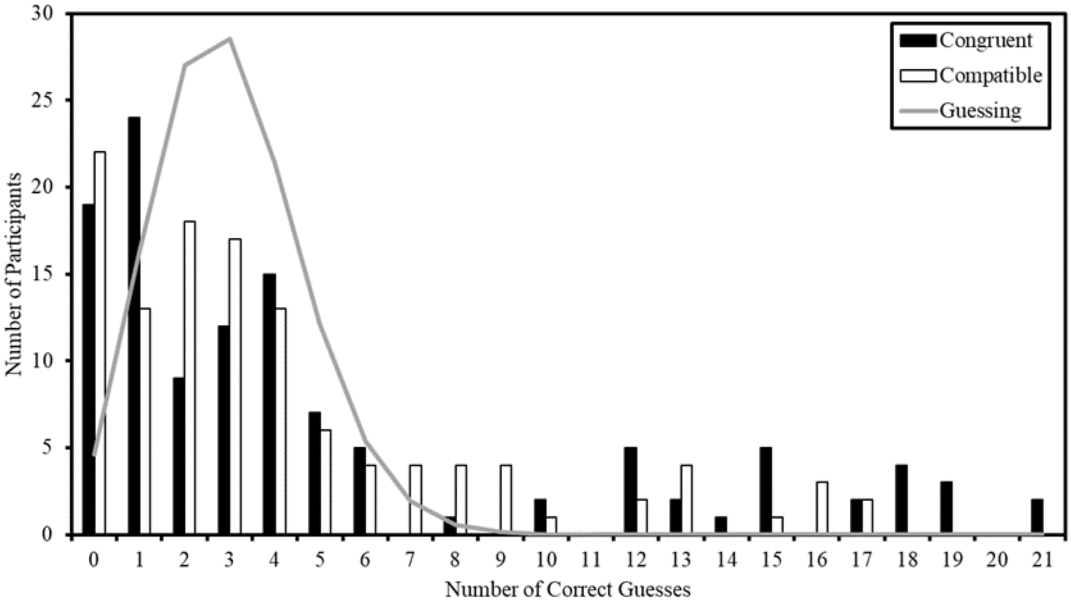
Error Rates

The *t*-tests analyses revealed no significant difference for error rates between congruent and control trials ($M_{congruent\text{-}control}$ = .730, *SD* = 4.82), *t*(117) = 1.645, *p* = .103, *d* = .151, BF_{10} = .304, or between compatible and control trials ($M_{compatible\text{-}control}$ = .307, *SD* = 4.01), *t*(117) = .831, *p* = .408, *d = .076*, BF_{10} = .164.

Test phase

The *t* tests on accuracy rates in the test phase (akin to the objective awareness phase in the previous experiments) revealed accuracy rates that were significantly above chance (1/7 or 14.3%) for both the congruent response (M = 24.7%, SD = 28.1), $t(117) = 4.00$, $p > .001$, d = .368, $BF_{10} > 100$, and the compatible response (M = 19.6%, SD = 20.4), $t(117) = 2.80$, $p = .003$, d = .258, $BF_{10} = 8.334$. Both of these effects, especially the congruency effect, seem to be due to a small number of outliers. Figure 13 shows the distribution of the results in the test phase. As can be seen, most of the participants seemed to be guessing (i.e., their results are under or slightly above chance guessing). However, few of them seemed to have enough preexisting knowledge about the congruency between note names and note positions, with some participants “guessing” 100% of the pairings correctly. Given that there was no way to *learn* the congruent pairings in the present experiment, this clearly indicates preexisting knowledge. The compatibility effect seems similar, but weaker, with an even smaller number of participants indicating the spatially compatible response well above chance.

Figure 13 - Experiment 3, distributions of the number of congruent and compatible guesses (out of 21) along with the expected number of correct responses if guessing alone.



Note. The guessing curve assumes that participants do not have a bias to repeat the same response to the same stimulus. The distribution would be flatter if participants have said bias, probably explaining the larger number of participants with a score near zero and multiples of three along with the smaller number of participants near the expected peak of the distribution.

Discussion

In Experiment 3, we tested for possible influences of the SMARC effect on the congruency effect we observed in Experiment 2. That is, we wanted to study whether the natural tendency of spatially coding the pitches could influence participants' responses in an incidental Stroop-like task. Our results did not show a significant difference in response times between spatially compatible and control trials. Interestingly, we also did not replicate the congruency effect in response times or errors despite a notably larger sample size. As already discussed, this absence of a congruency effect is actually consistent with a number of prior reports with a similar or (in some cases) near identical procedure. The significant congruency effect observed in Experiment 2 may therefore have been a Type 1 error.

On the other hand, the nonmusicians responded significantly above chance in the test phase with the congruent response. Given that there was no way for participants to learn the congruent pairings without the current experiment, this clearly indicates that some small number of participants *did* have preexisting sight-reading knowledge. The same test phase also revealed elevated numbers of spatially compatible responses. These latter results may suggest that the natural inclination for spatially coding pitches can influence performance in some cases, such as in a more explicit judgement task.

General Discussion

In our study, we were interested in investigating early acquisition of sight-reading skills in an incidental learning procedure. That is, can nonmusicians with no prior familiarity with music reading rapidly acquire knowledge of standard notation that in turn produces automatic influences on performance in a similar way to that observed in skilled musicians? As hypothesized, despite a very short learning phase (336 trials, approximately 15 min) and slightly more complex material than those used in previous incidental learning procedures (e.g., words and colors), nonmusicians produced a robust contingency effect during the learning and subsequent test phases in both the deliberate (Experiment 1) and incidental (Experiments 1 and 2) learning groups.

Musicians can easily read music symbols and Grégoire et al. (2013) pointed out that the Musical Stroop Effect can be explained by the automaticity of the learned association between the note position and note name. Musicians cannot avoid “naming” the note-position just as skilled readers cannot avoid reading color-words in the regular Stroop task. Furthermore, Schön et al. (2001, 2002) proposed that musicians rely on different types of translation when reading music. For instance, playing-like (i.e., visuomotor translation) and naming-like (visual-verbal translation) transcodings are important to automatize the process of sight-reading. In general, sight-reading seems to be a complex process based on visuomotor integration (Gudmundsdottir, 2010).

In the present report, we showed that recently acquired associations, even if only learned incidentally, can produce the same automatic influences on behavior. Although our predictive stimulus (note-position) was not task-relevant (i.e., not the target stimulus), it produced an effect on performance, anyway. That is, our participants were able to learn the associations between note names and note positions as well as the corresponding actions. As mentioned in the Introduction, it may be the case that learning the contingencies between the predictive stimulus and the target drives the prediction of the motor response. Further, it is likely the case that learning is so rapid because participants can gain extensive practice of the stimulus-response pairings in a short period of time, which often is not the case with more deliberate learning procedures (Logan & Klapp, 1991). Although contingency learning has been observed in numerous learning paradigms (e.g., the color-word contingency learning paradigm), here we show for the first time the presence of the contingency effect in a music-related task. We were able to prove that the same sort of learning observed between simple stimulus pairs (e.g., colors and words) is also observable with more complex (e.g., in terms of the number of stimuli presented and the number of associations to learn) and more ecological musical materials.

The main aim of Experiment 2 was to investigate to which extent previous knowledge can influence the contingency effect found in Experiment 1. We asked for nonmusician participants who do not know how to sight read, though there is always a risk that participants have studied music at school and remember more than they imagined. We therefore scrambled the note-name to note-position correspondences. Reassuringly, a contingency effect was still found in Experiment 2, suggesting once again a rapid incidental learning of the presented associations. As the associations between note positions and responses to the note names were *not* congruent in Experiment 2, this learning effect could not have been due to preexisting sight-reading knowledge. However, in Experiment 2 a congruency effect was also found, suggesting the presence of previous musical knowledge in some participants, possibly due to music training at school. Based on this evidence, it is possible that the contingency effect in Experiment 1 was influenced by the congruency effect (i.e., because in Experiment 1, contingency was confounded with congruency, since all high-contingency trials were congruent and all low-contingency trials incongruent). In any case, our results, though indicating that undisclosed musical knowledge might impact the measure of learning if only congruent associations are used, true contingency learning is still present during the learning phase.

To further elucidate the congruency effect observed in Experiment 2, we ran a third experiment in which we investigated the hypothesis that the congruency effect in Experiment 2 was influenced by the SMARC effect. As previously mentioned, the SMARC effect refers to the natural human tendency for spatially coding pitches (Rusconi et al., 2006), even without the presence of an actual sound (Ariga & Saito, 2019). Based on this premise, in Experiment 3 we dissociated congruency from stimulus-response spatial compatibility. In this way, we distinguished between congruent trials, in which the note name was congruent with the note position (e.g., the note name "do" in the position for "do") and compatible trials, where the spatial position of the target was compatible to the spatial position of the response key on the keyboard (e.g., the note name "do" in the bottommost "fa" note position when the key responses were ordered from "do" to "si"). Our aim was to measure to which degree the previously observed congruency effect was due to preexisting sight-reading knowledge (as measured by congruency) and/or to a SMARC-like spatial compatibility effect. In response times and errors, we failed to replicate the finding of Experiment 2, with no congruency or compatibility effect. Potentially, this might indicate that the significant congruency effect in Experiment 2 was due to Type 1 error, or that some other seemingly trivial difference between Experiments 2 and 3 was responsible for the different outcomes. However, while in our study participants were engaged in an incidental learning procedure, previous SMARC studies (Ariga & Saito, 2019; Rusconi et al., 2006) asked participants for explicit judgements. It is worth noting that in our study also, when nonmusicians were required to provide an explicit response in the test phase, their performance was significantly above chance level, suggesting the presence of a SMARC effect. We also observed above-chance congruent responses in the same test phase, clearly indicating that some small number of participants did have some preexisting knowledge. This suggests that future studies that aim for a "pure" measure of learning might be best adapted with some form of pretest of preexisting knowledge and/or nonspatially compatible stimulus-response mappings.

Additionally, as previously hypothesized, overall test phase accuracy (objective awareness), in both Experiments 1 and 2, indicates that nonmusicians performed above chance, suggesting that they were able to learn the associations that they were exposed to and even to verbalize this knowledge. However, in Experiment 1, a significant difference in favor of the deliberate learning group in the objective awareness results suggests that deliberate learning boosts learning more than purely incidental learning does. Previous research showed that to learn contingencies, being attentive to the predictive dimension is important (Eitam et al., 2009; Jiang & Chun, 2001). If this is the case for the deliberate learning group, then it is not surprising that they gave more accurate responses in the objective awareness phase than the incidental learning group did. At the same time, the evidence in favor of the deliberate learning group may simply suggest that learning in a deliberate way might aid more during explicit reporting (in the objective awareness phase) than in the case of automatic execution (in the learning phase). In other words, our objective awareness phase specifically required participants to express an explicit judgment, unlike the learning phase where participants were asked for automatic execution. Although, the deliberate learning group reported more accurate response in the objective awareness phase than the incidental one, the nonsignificant Contingency × Group interaction in the learning phase suggests that the two groups were able to automatize the learned contingencies in a quite similar way. Thus, deliberate learning may provide an advantage when it comes to explicit reporting, but perhaps may not confer the same advantage for automatization of contingency knowledge.

Although the observed acquisition of sight-reading knowledge may seem implausibly fast to some readers, such results are not a surprise when considering prior contingency learning work with other stimuli. As previously mentioned in the Introduction, contingency learning paradigms like the present one allow for extremely rapid acquisition of the associations between stimuli in a task (Lin & MacLeod, 2018; Schmidt et al., 2010; Schmidt & De Houwer, 2016), therefore the present results are completely coherent with past work using related, nonmusical learning procedures.

We note that our aim was not to claim that a procedure such as ours can replace other types of deliberate practice, which are more goal-oriented (Ericsson et al., 1993; Ericsson & Harwell, 2019; Mishra, 2014). On the contrary, we believe that the acquisition of complex skills, such as sight-reading, can benefit from both deliberate and incidental learning procedures. On one side, more deliberate training can guide the acquisition of instrument-specific skills, such as effortful strategies to improve the technical movements of the bow on the strings to play the violin. On the other hand, an incidental learning procedure such as that used in the present report can help with the automatization of visuomotor integration, favoring sight-reading performance.

As one potential limitation, in the current study participants responded to note-names and learned about the note-positions incidentally. We did this for a few reasons. Most importantly, the current methodology allowed us to study the automatic (i.e., stimulus-driven) influences of note-position knowledge on performance (e.g., akin to the musical Stroop with experienced musicians or the influence of color words on color naming in the traditional Stroop paradigm). Learning may, however, be even stronger and faster if participants respond to the note positions directly (i.e., the note-position is the target, rather than the task-irrelevant but informative stimulus). We are currently investigating this in an ongoing study. Furthermore, as already noted in the Introduction, we used an imperfect contingency manipulation (75% high-contingency vs. 25% low-contingency). Although this was done to measure learning while it was occurring, a perfect contingency manipulation (e.g., using a 100% congruency between note-names and note-positions) may further help learning, especially in a real-world application (e.g., helping nonmusicians to acquire sight-reading skills with a learning app). This point is the object of another ongoing study we are currently conducting.

As another limitation, although we used arbitrary stimulus-key assignments similar to the ones that musicians practice on their instrument (especially piano), we did not use real instruments for learning. In future research, using the same logic of this study, it may be interesting to use a very similar piano response modality, or also other types of instruments (like string or wind instruments). A vocal response modality (e.g., singing) could also be used. Globally, the goal was to show that this type of position-to-action learning can occur rapidly with an appropriately designed learning procedure, but real-world applications to actual instruments remains to be explored. Furthermore, although here we mostly focused on the acquisition and automatization of the associations between spatial positions and motoric responses, previous research suggested that auditory stimuli are important to train sight-reading skills. That is, sight-reading benefits greatly from an integration of visual, auditory, and motor components (Brodsky et al., 2003, 2008; Gromko, 2004; Hayward & Eastlund Gromko, 2009), rather than just visuomotor integration (Gudmundsdottir, 2010). In other words, learning what the note positions sound like can facilitate sight-reading skills. In on-going studies, we are investigating the role of auditory stimuli in learning in our task, to further test the facilitative benefit of auditory stimuli in the acquisition of sight-reading skills.

In conclusion, we showed the presence of the contingency effect in an incidental music contingency procedure, as well as the ability to verbalize the knowledge that was incidentally (or deliberately) acquired. Such findings are exciting, because they suggest that a seemingly difficult-to-learn music skill, sight-reading, can be learned much more quickly and easily than previously assumed. In the short-term, we hope that this paper will serve as the starting point for further investigations of the incidental learning of complex material, musical or otherwise, including investigations of ways to reinforce learning even further. In the long-term, this study may open up a new line of research to implement the same or similar approaches in an applied setting to help novices (whether in a musical and nonmusical context) to acquire valued skills with greater ease.

Chapter 3 – learning and pitch identification

Incidentally acquiring pitch-label associations with a musical contingency learning task

Claudia Iorio and James R. Schmidt

Introduction

Most humans, even nonmusicians, 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. *Implicit learning* occurs without the intentional goal to learn and without conscious awareness of what has been learned (Cleeremans et al., 1998; Reber, 1989). The present work focuses on *incidental learning*, which is learning that occurs without the explicit intention to learn (Kerka, 2000), but that may or may not be unconscious. The implicit learning of music material has been already investigated in prior work, such as the implicit 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). However, the relationship between implicit or incidental learning and the internalization of pitch identities has yet to be studied. Therefore, here we wanted to investigate the potential utility of an incidental learning procedure to help nonmusicians and musicians to detect pitches by ear (i.e., to acquire or improve pitch-label associations).

Pitch detection and absolute pitch

When seeing a color (e.g., yellow), everyone with normal color vision is able to identify and name the color easily. It is certainly not necessary to compare, for instance, yellow and blue to determine the current color (yellow). Interestingly, the same cannot be said for pitch detection. *Absolute pitch* (AP) is defined as the ability to automatically and effortless identify and name pitches without any external reference (Deutsch, 2013a; Levitin, 1994; Levitin & Rogers, 2005; Takeuchi & Hulse, 1993). AP is considered to be a rare ability and, indeed, it is only present in a small percentage of the population (Miyazaki et al., 2012; Takeuchi & Hulse, 1993; Ward, 1999), even among skilled musicians. Instead, most musicians rely on relative pitch (RP) when identifying the name of the pitches. That is, RP possessors are able to identify and name pitches using a comparison-based strategy in which they name the pitch through the processing of the melodic and harmonic relations between tones (Levitin, 1994; Levitin & Rogers, 2005; Takeuchi & Hulse, 1993). For instance, after hearing and being given the identity of one tone (e.g., do), a RP possessor can determine the identity of a new tone (e.g., mi) by detecting the interval (i.e., number of semitones, akin the smallest distance between two notes) between the first known note and the new one. An AP possessor does not need to make such a comparison and perceives the note identity directly.

The reason why very few people possess AP is still not completely clear, especially for musicians who spend years playing and hearing notes. Some authors proposed that there is a strong genetic component, which is supported by many different studies (Athos et al., 2007; Baharloo et al., 2000; Deutsch, 2013a; Theusch & Gitschier, 2011). For instance, shared AP between siblings is much more common in identical than fraternal twins (Theusch & Gitschier, 2011), it appears at a very young age (Deutsch, 2013a), and AP possessors show a unique structured brain circuity (Bermudez & Zatorre, 2009; Loui et al., 2011; Schulze et al., 2009). In addition, according to the critical period hypothesis, an early onset of music training (e.g., before 4 or 5 years old) can benefit the development of AP (Crozier, 1997; Deutsch, 2013b; Deutsch et al., 2006; Miyazaki & Ogawa, 2006), with AP rarely observed for those who started learning music at a later age.

Interestingly, although naming pitches in an absolute way is a rare ability, most people appear to have some degree of what is sometimes termed *implicit AP*. As an illustration, people that are not able to name a pitch without an external reference 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. Similarly, they can correctly reproduce familiar melodies with a reasonable degree of accuracy (Levitin, 1994). Levitin (1994; see also Levitin & Rogers, 2005) suggested a distinction between different traits of AP, namely pitch memory and pitch labelling. The former is proposed to be the widespread ability to store in a long-term memory system information about pitches (i.e., in an absolute way). Pitch labelling, on the other hand, is specific for AP possessors and it refers to the ability to perfectly label the pitches. While both AP and RP possessors can store pitch information in a stable long-term form, only AP possessors immediately retrieve the pitch's name from a "pitch template" that they have stored in long-term memory. On the contrary, RP possessors process the melodic and harmonic relations to compare pitch information to an "interval template" stored in long-term memory. In other words, AP possessors are able to automatically retrieve pitch labels from memory, while RP possessors require a strategic comparison process (Levitin & Rogers, 2005). Therefore, AP possessors are usually not only more accurate but also faster when labelling pitches compared to RP possessors. Specifically, the accuracy of AP possessors is well above chance (between 50% and 100%) when tested for AP (Levitin & Rogers, 2005, Miyazaki 1988), and their identification times are rapid (between 1.5s and 3s), unlike for RP possessors (Bermudez & Zatorre, 2009; Miyazaki, 1990; Takeuchi & Hulse, 1993; Van Hedger et al., 2019; Wong, Lui, et al., 2020). These results suggested that AP possessors are able to automatically retrieve pitch-label associations, unlike those relying on the interval comparison strategy used by RP possessors.

Neuroimaging studies revealed that when labelling pitches AP possessors showed an activation in the posterior dorsolateral frontal cortex, an area usually associated to conditional associative learning. On the contrary, RP possessors showed an activation in areas associated to working memory (Zatorre et al., 1998). Specifically, while RP possessors seem to rely more on working memory resources while identifying pitches, AP possessors retrieve the pitch-label association directly from long term memory (Hirose et al., 2002; Shahin et al., 2003; Zatorre, 2003; Zatorre et al., 1998).

However, 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 to learn AP. They demonstrated that after explicit and effortful training, adults (musicians and nonmusicians) were able to improve pitch detection. Some showed performance at posttest similar to AP possessors (though their pretest scores were already reasonably good). It is not necessarily controversial to suggest that pitch detection 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 detection abilities at all could learn to have AP beyond some of the stricter criteria.

In the present work, we explore a novel approach to improving pitch detection. Our goals are notably different than the preexisting 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 explore an entirely new approach to training pitch detection skills and to determine whether nonmusicians with no pitch detection abilities (and musicians in Experiment 2) can easily and rapidly learn pitch-label associations, improving their pitch detection. As discussed in more detail below, we use a short-duration incidental learning procedure, with the hypothesis that this sort of training will produce rapid early learning. We do not aim to argue that our participants develop AP as typically defined, though this may be a question for future research (as discussed in more detail in the General Discussion).

Current work: the music contingency learning procedure

In the present work, we make use of an incidental learning task. As mentioned above, incidental learning refers to the process of acquiring new information without the goal to learn

(Kerka, 2000). That is, the participant is instructed to do one thing (e.g., identify a target stimulus) and learning of regularities that exist in the task (e.g., that another task-irrelevant stimulus is correlated with and thus “predictive” of the target stimulus) is “incidental” to the explicit task goal. This contrasts with intentional or deliberate learning, where the instructed goal of the task is to try to learn the regularities. Note that a related distinction is made between implicit and explicit learning. Implicit learning (Reber, 1989) is considered to be both incidental and unconscious. There is considerable debate about the conscious vs. nonconscious nature of learning (Cleeremans et al., 1998). Although we do take some awareness measures, it is not our goal to make any hard claims about whether any learning observed is due to conscious or unconscious knowledge of the regularities. As such, we refer to our task as an incidental learning procedure below.

Our particular training approach makes use of a *contingency learning* procedure. 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 a typical contingency learning procedure (e.g., the color-word contingency learning procedure of Schmidt et al., 2007) participants are exposed to regularities between the target (e.g., color) and a second nontarget stimulus (e.g., word). For example, the word “move” might be presented frequently in blue, but rarely in green or red. Although participants are not instructed to learn the associations between the stimuli, they do learn them. This learning is extremely rapid: 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* (Schmidt et al., 2007; for related learning procedures, Carlson & Flowers, 1996; Miller, 1987; Mordkoff & Halterman, 2008; Musen & Squire, 1993). Similarly, fast learning is observed in a range of other incidental learning procedures (e.g., sequence learning, Nissen & Bullemer, 1987; Turk-Browne et al., 2005; 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; hidden covariation detection, Lewicki, 1985, 1986; Lewicki et al., 1992).

Not only are we naturally able to detect regularities between events (Kelly & Martin, 1994; Zacks & Hasher, 2002), but we are also able to learn and use this information in a variety of useful contexts (e.g., in language acquisition; Aslin et al., 1998; Saffran, Aslin, et al., 1996; Saffran et al., 1997; Saffran, Newport, et al., 1996). In a series of studies (Iorio et al., 2022, Schmidt et al. 2022), we applied this logic to another musical skill: sight-reading. Using a musical contingency learning procedure, we showed that nonmusicians were able to learn the associations between note positions and note names and use them to correctly identify the name of the note position presented on the musical staff in a note naming task. Specifically, 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., the note name “do” was presented much more often with the congruent position “do” than with any other incongruent positions). 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, nonmusician participants responded robustly faster to congruent than to incongruent trials during the learning phase. This indicates 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. As mentioned above, automaticity like this is relevant for pitch detection, as well.

Similar to visual Stroop tasks (see Grégoire et al. for a musical version, 2013; see Schmidt et al., 2007 for the color-word paradigm), 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). 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 and Henik (2012) compared performance in a tone naming task and note naming task between AP and non AP 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. It was our hypothesis that participants can not only be trained to improve accuracy in pitch detection 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., 2022) to measure the automaticity of pitch processing in nonmusicians and musicians. In our previous study, we aimed to study the acquisition of note name/note position associations in nonmusicians. Therefore, participants were asked to respond to the note name (target) that appeared in one specific position (cue) on the music staff. The note name could be congruent with the note position (e.g., the name "do" in the note position for "do") or incongruent (e.g., the name "mi" in the note position for "do"). Here, our aim was to study the acquisition of pitch-label associations, therefore we replaced the note positions with tones. Specifically, participants heard a tone before the presentation of each note name. Each tone was presented most often with the congruent note name and rarely with the incongruent note names, as illustrated in Table 2.

Table 2 - Contingency manipulation.

Note Name	Note Position						
	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 note position "fa" is presented much more often with the note name "fa" (high contingency trials) than with the other note names (low contingency trials).

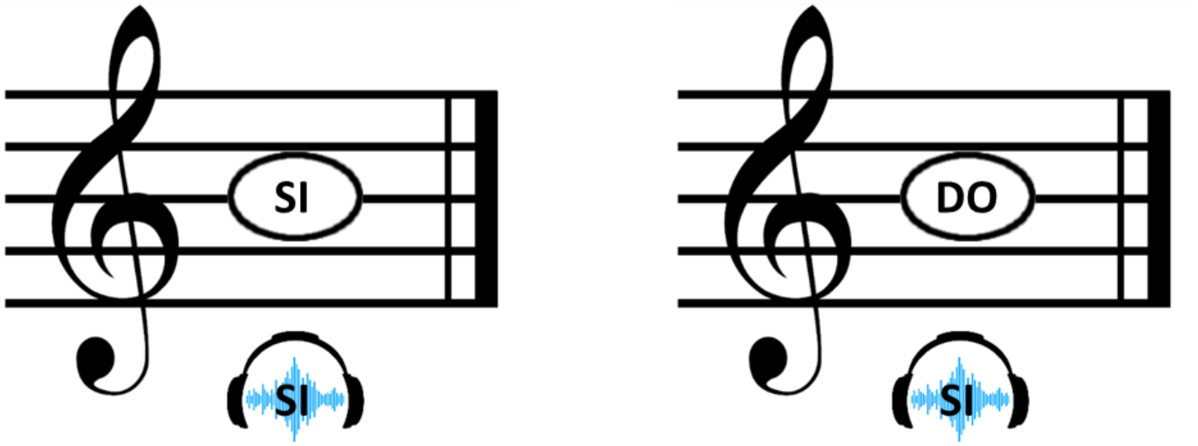
Our key hypothesis, across the three experiments to follow, is that participants will be able to learn (or improve) their pitch detection abilities. This should be reflected both in an increase in explicit identification of note pitches after training, and more automatic effects on performance (i.e., faster and more accurate responses to high-contingency congruent trials relative to low-contingency incongruent trials). More specifically, in Experiment 1 we investigated whether nonmusicians are able to learn pitch-label associations in our incidental musical learning procedure. 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. As a supplementary question, we also asked whether participants might benefit from a multiple-cues presentation, with both note positions and tones presented as cues, compared to a group presented only the auditory tone cues. In Experiment 2, we investigated whether a significant improvement in labelling pitches can also be seen in musicians (i.e., enhance their pitch identification performance). In Experiment 3, we investigated whether the observed learning effects persist in memory. Specifically, one week after the end of the learning phase, nonmusician participants were tested to see whether they were still able to identify and name the pitches. As an added question, we also assessed intentionality in Experiment 3. Previous results indicate that being aware of the contingencies can "boost" learning (Schmidt & De Houwer, 2012a, 2012b; for a similar result in sequence learning, see Destrebecqz, 2004). Therefore, to assess the role of intentionality in memory consolidation, in Experiment 3 nonmusicians were divided into a deliberate learning group, instructed to pay attention to the contingencies, and an incidental learning group, who received no instructions about the presence of the contingencies.

To summarize our predictions, in Experiment 1 we hypothesized that nonmusician participants would learn the associations quickly, showing both improved accuracy in explicit pitch detection and automatic effects on performance during the learning phase (i.e., faster and more accurate responses to congruent trials). Concerning the multiple-cues group (who were presented also with note positions), we considered two alternative hypotheses. One possibility is that performance will be improved relative to the tone-cue group, as the note positions might provide an added visual cue to encode associations between tones and note names. Alternatively, overshadowing might occur. Specifically, the presence of associations between note positions and note names might impair the learning of associations between tones and note names. For Experiment 2, we hypothesized that our contingency learning procedure would help musicians to improve their pitch identification performance. Finally, in Experiment 3 we had two main hypotheses. First, we hypothesized that the deliberate learning group would show even better performance in pitch identification than the incidental learning group. Second, we anticipated that contingency knowledge would persist in memory and that follow-up recall test performance would be well above chance guessing.

Experiment 1

In Experiment 1, we wanted to investigate whether nonmusicians were able to incidentally learn pitch-label associations. For this purpose, we used a modified version of the musical contingency learning procedure from our previous studies (Iorio et al., 2022, Schmidt et al. 2022), 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), we 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.

Figure 14 - Example of high and low trials.



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

Our primary hypothesis is that both groups of nonmusician participants will incidentally learn the pitch-label associations. In particular, we anticipate faster (and possibly more accurate) responses to congruent (high-contingency) trials than to incongruent (low-contingency) trials (see Figure 14 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 (Kamin, 1969).

Method

Participants

119 participants, recruited online on 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. Sixteen participants reported to have AP. 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 AP. 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 AP: $t(117) = .044$, $p = .946$, $d = .012$, $BF_{10} = .271$, $M_{\text{absolute pitch participants}} = 16.7\%$, $SD = 9.52$, highest score = 33.33%; $M_{\text{remaining participants}} = 16.8\%$, $SD = 14.0$ highest score = 85.71%. Therefore, we did not exclude these participants from the analysis. However, although three participants in the tone-cue group declared not to have AP, their performance in the pretest were between 60% and 80% , 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.

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 Psytoolkit, a web-based software that allows reliable RTs as shown from previous research (Stoet, 2010, 2016), also with music stimuli (Armitage & Eerola, 2020). The auditory stimuli 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. 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, 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, 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.

Participants were randomly assigned to one of the two groups. In both groups, participants were asked to respond to the note names (target) in the main learning phase. However, 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.

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 absolute 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?"*

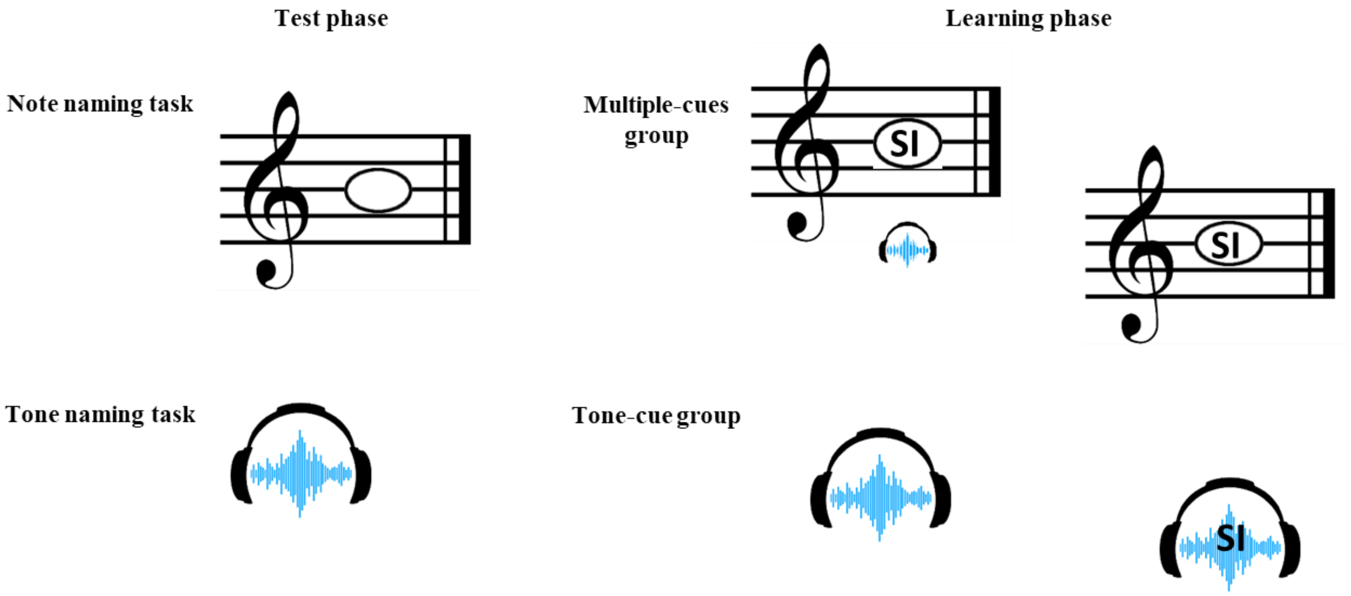
This question was primarily used for screening purposes, along with the pretest scores, as described above in the Participants section.

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).

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 interested 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 detection (or sight-reading) skills in the absence of music training, but the pretests allowed us to both (a) screen for undisclosed preexisting knowledge and (b) to establish a control for pre/post improvement scores.

The same procedure was added at the end of the experiment (i.e., after the learning phase, to be describe next) as a posttest. This allowed us to compare participants' performance before and after the learning process. However, as previously mentioned, 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 15), in which they had to guess the name of the tone (no limit time; 21 trials). However, the note-position naming task (Figure 15), 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.

Figure 15 - Schematic description of the learning and the test phases.



Note. On the left panel of the image an example of the tone and note position naming tasks. On the top right, an example of how the learning phase looked like for the multiple-cues group. On the bottom right is the learning phase for the tone-cue group.

Immediately after the pretest phase, participants started the learning phase that differed between the groups as shown in Figure 15. 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 incorrectly 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., 2022) congruent pairings (e.g., the tone "fa" for the note name "fa"; high-contingency trials) and 10% incongruent trails (e.g., the tone "fa" for the note name "do"; low-contingency trials; see Table 3 for the proportion of the contingency manipulation. The congruency (or contingency learning) effect was measured as the difference between low- and high- contingency trials.

Table 3 - 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).

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. 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

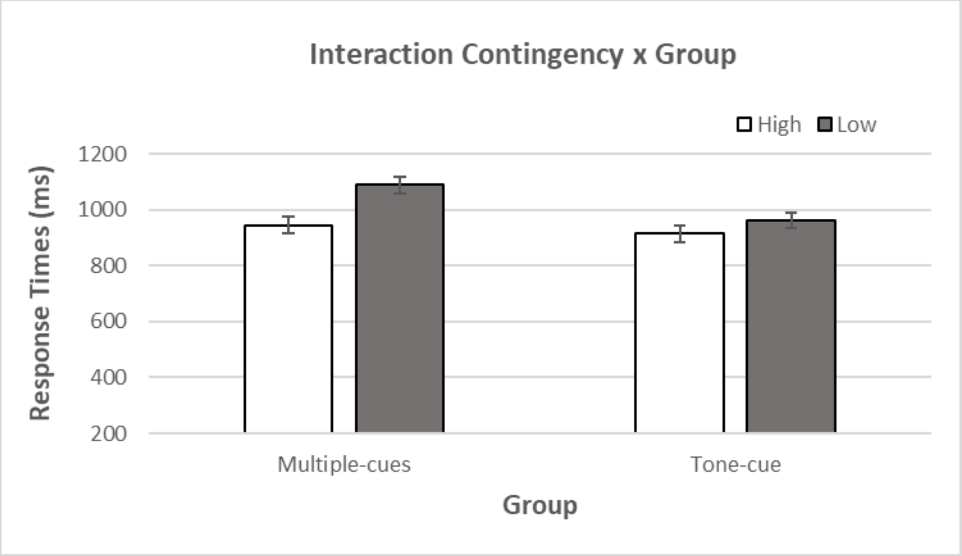
We conducted analyses on the learning and the test phases. For the learning phases, we conducted a repeated measures ANOVA on correct RTs and error rates to assess the overall main effects of contingency, group, and the interaction between them. Trials in which participants failed to respond in 3000 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%). All analyses were evaluated at the $\alpha = .05$ level of significance. Additionally, 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 reported the Bayes factor BF_{10}, with values between 3 and 10 supporting a 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 16. 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 ($M = 935$ ms, $SD = 231$) than for low-contingency trials ($M = 1027$ 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 faster response in the tone-cue group ($M_{high\_trials} = 925$ ms, $SD = 202$; $M_{low\_trials} = 962$ ms, $SD = 199$), than in the multiple-cues group ($M_{high\_trials} = 946$ ms, $SD = 257$; $M_{low\_trials} = 1089$ ms, $SD = 232$). 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, $M_{\text{low-high}} = 37.6$, $SD = 78.0$; $t(56) = 3.64$, $p < .001$, $d = .482$, $BF_{10} > 100$.

Figure 16 - Interaction between contingency effect and Group in Experiment 1.

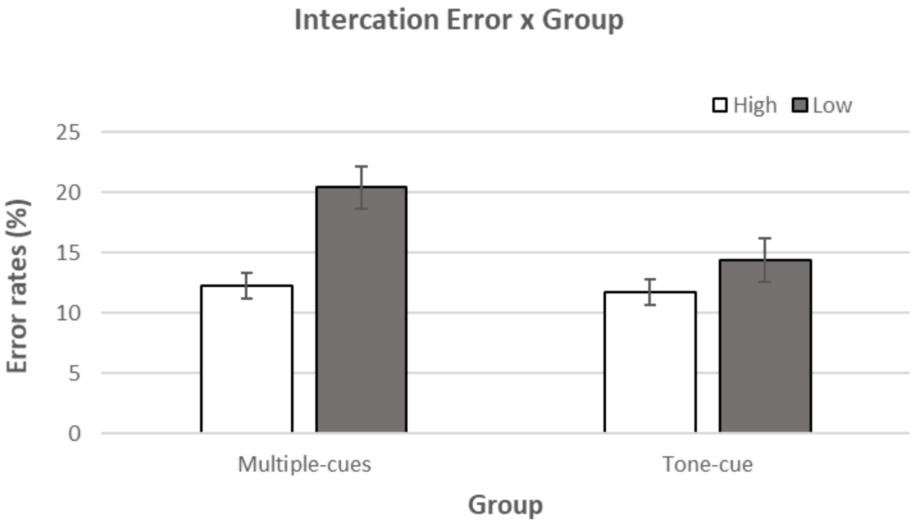


Note. Interaction between Contingency (High and Low) and Group (multiple-cues vs. tone-cue), standard error bars are shown in the figure.

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 nonsignificant 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 17, $F(1,114) = 12.5$, $p <.001$, $\eta^2 = .099$, $BF_{10} = 43.84$ (multiple-cue group: $M_{high} = 12.3\%$, $SD = 8.24\%$, $M_{low} = 20.4\%$, $SD = 13.2\%$; tone-cue group: $M_{high} = 12.0\%$, $SD = 8.23\%$, $M_{low} = 13.9\%$, $SD = 13.4\%$) . 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 $M_{\text{low-high}} = 1.96\%$, $SD = 5.16\%$; $t(56) = 1.78$, $p = .081$, $d = .236$, $BF_{10} = .631$.

Figure 17 - 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.

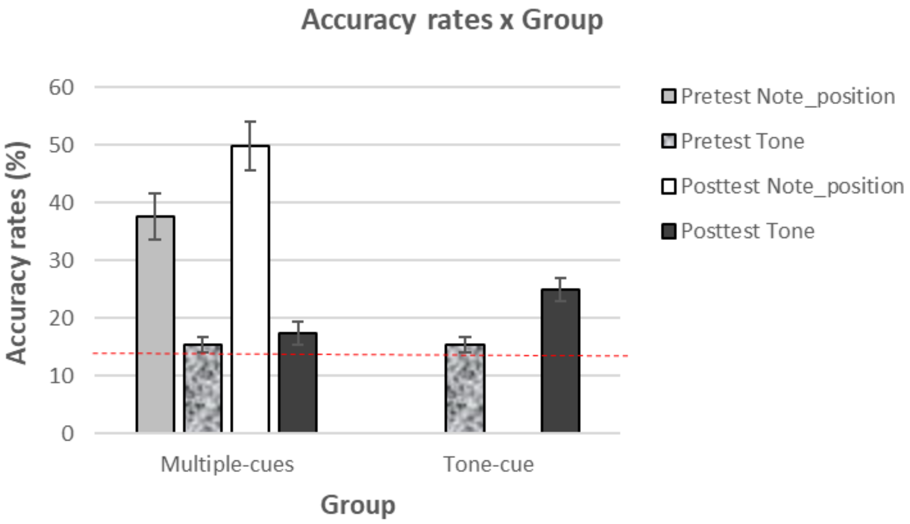
Pre/Posttest phases

Since only the multiple-cues group saw note positions as cues in the learning phase, the subjective awareness question for note positions is calculated on the total number of participants in the multiple-cues group. Accordingly, 60.40% (38 of 59) participants became aware of the contingencies between note names and note positions. On the contrary, the percentage of participants aware of the contingencies between tones and note names is computed on the total number of participants in both the groups and it indicated that 61,34% (73 of 119) noticed the contingencies between note names and tones.

The *t*-tests for pretest and posttest accuracy, as shown in Figure 18, showed that in the note position naming task, 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, $t(58) = 8.32$, $p < .001$, $d = 1.08$, $BF_{10} > 100$, $M = 49.8\%$, $SD = 32.8\%$ (significantly higher performance were reported in the posttest compared to pretest, $t(58) = 3.46$, $p = .001$, $d = .450$, $BF_{10} = 26.5$). In the tone naming task, 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 reported performance not significantly above chance in the pretest $t(56) = .814$, $p = .419$, $d = .108$, $BF_{10} = .198$, *M* = 15.4%, *SD* = 9.94% and significantly above chance in the posttest $t(56) = 4.94$, $p > .001$, $d = .654$, $BF_{10} > 100$, *M* = 25.0%, *SD* = 16.33%; the data showed a significant improvement between the pretest and posttest $t(56) = 4.33$, $p > .001$, $d = .574$, $BF_{10} > 100$.

Most importantly, we also ran an ANOVA with the factors Test (pret vs. post) and Group (multiple-cues vs. tone-cue) on the tone naming accuracy rates to analyze for possible interactions. The results showed a main significant 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 18.

Figure 18 - 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) in the note position naming task ($M_{pretest}$ = 37.6%, *SD* = 30.1%; $M_{posttest}$ = 49.8%, *SD* = 32.8%) and the tone naming task (multiple-cues: $M_{pretest}$ = 15.4%, *SD* = 10.4%, $M_{posttest}$ = 17.4%, *SD* = 12.7%; tone-cue: $M_{pretest}$ = 15.4%, *SD* = 9.94%, $M_{posttest}$ = 25.0%, *SD* = 16.33%). Standard error bars and accuracy chance guessing at 14.3% (in red) are shown in the figure.

Overall, these results indicate that both groups of participants were able to learn pitch-label associations. However, 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.

Discussion

In Experiment 1, we wanted to study whether nonmuscians 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 and were able to respond above chance in the tests phases in line with previous findings in the contingency learning literature (Schmidt & De Houwer, 2019, Iorio et al., 2022). However, overall, 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 subskills (Mishra, 2014a), 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, showing an overshadowing effect.

Experiment 2

In Experiment 1, we studied the more experimentally “pure” case of nonmusicians learning to identify note pitches. This sort of sample would also correspond to novice musicians just beginning to learn music. Learning to improve pitch detection 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. 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. Our primary hypothesis for the present experiment was that musicians would be able to improve their pitch detection, 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 detection 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 AP. 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 from the participants recruited online. 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). In Experiment 1, the lowest note name "fa" corresponded to the most left key on the keyboard "Z". Possibly, this could help with the acquisition of key-label responses based on the research about the SMARC effect. As a small note, we did not find a facilitation effect of the SMARC effect on the acquisition of the note name/note position association in our previous work (Iorio et al., 2022). However, to control for possible influence of the SMARC effect in our paradigm, we compared performance in two groups. 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 refers to "do" to "si" note names. In this group, there was no spatial compatibility between the tones and the responses (e.g., the leftmost key "D" 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

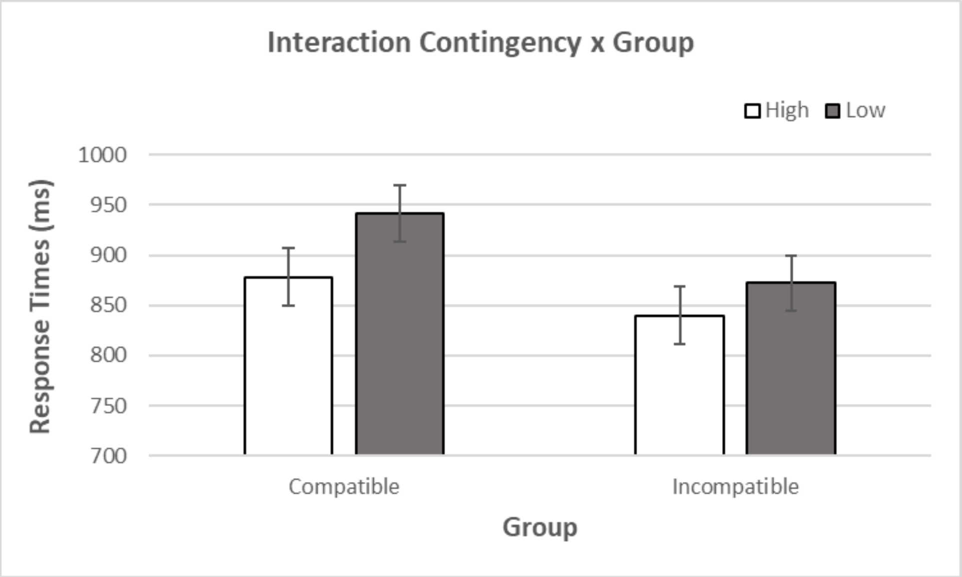
As in Experiment 1, we ran an ANOVA on RTs and error rates for the learning phase and *t*-tests on accuracy for the test phases. However, here we additionally added information about the start of musical training as a covariate in our analysis. 12.77% of the trials on the total number of participants were eliminated based on the same criteria used in Experiment1 (i.e., trials in which participants failed to respond in 3000 ms). All analyses were evaluated at the α = .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 ms, SD = 209) than for low-contingency trials (M = 909 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 19), $F(1,98) = 4.12$, $p = .045$, $\eta^2 = .040$, BF_{10} = 1.02 (compatible group: $\mathrm{M}_{\text{high-contingency}}$= 878 ms, SD = 214, $\mathrm{M}_{\text{low-contingency}}$ = 942 ms, SD = 227; incompatible group: $\mathrm{M}_{\text{high-contingency}}$= 840 ms, SD = 204, $\mathrm{M}_{\text{low-contingency}}$ = 872 ms, SD = 210). 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; $t(47) = 3.21$, $p = .002$, $d = .436$, $BF_{10} = 13.1$. The interaction between Contingency and beginning of the musical training was not significant $F(1,98) = .621$, $p = .433$, $\eta^2 = .006$, $BF_{10} = .301$.

Figure 19 - 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.

Error Rates

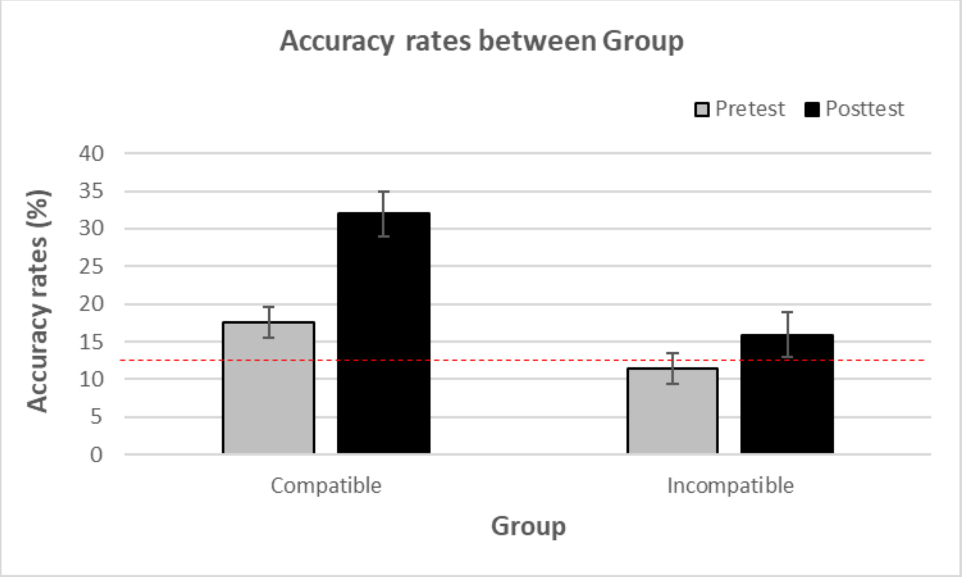
The repeated measures ANOVA for errors with the factors Contingency (high vs. low) and Group (compatible vs. incompatible) revealed a main effect of Contingency, $F(1,99) = 17.51$, $p < .001$, $\eta^2 = .150$, $BF_{10} > 100$ (more errors for low trials $M = 15.3\%$, $SD = 11.8$, compared to high trials $M = 12.5\%$, $SD = 9.75$), and a nonsignificant 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% (77 of 108) participants noticed the contingencies between the tones and the note names. *T*-tests for pretest and posttest accuracy rates, shown in Figure 20, 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, and 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 significant 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.

Figure 20 - 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.

Furthermore, we ran a repeated measures ANOVA with the factors Test (pre vs. post) and Group (compatible vs. incompatible) on the tone naming accuracy to analyze for possible interactions. The main effect of Test was significant *F*(1,99) = 27.80, *p* < .001, η^2 = .219, BF_{10} > 100 as well as the main effect of Group *F*(1,99) = 17.4, *p* < .001, η^2 = .150, BF_{10} > 100. Also the interaction Test × Group was significant *F*(1,99) = 7.70, *p* = .007,

η^2 = .072, BF_{10} = 5.98 (pre-test: $M_{compatible}$ = 17.6%, *SD*= 14.5; $M_{incompatible}$ = 11.4% *SD*=10.1. post-test: $M_{compatible}$ = 32.0%, *SD*= 23.1; $M_{incompatible}$ = 15.9% *SD*=13.0

Discussion

In the above 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 in the test phases revealed that only the compatible group performed significantly above chance in the posttest with 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 showed in previous research (Ariga & Saito, 2019; Rusconi et al., 2006). Participants can learn the contingencies in either case, but spatial compatibility helps.

Experiment 3

In Experiment 3, we extend the results of the preceding experiments in two ways. First, 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, 2012a, 2012d). 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. This instruction effect is not always robust. For instance, in the above-mentioned sight-reading studies (Iorio et al., 2022) we did not find any significant differences in the learning phase between participants that were aware of the contingencies and the one 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. Second, we tested whether pitch learning persists over time. In particular, a second posttest phase was conducted approximately one week after the initial learning experiment. We hypothesized that posttest scores would still be increased after the delay (i.e., that the learned pitch information remains in memory).

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 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 he reported accuracy rates 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 in 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 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.[1]

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, here we discarded 13.37% of the data.

---

[1] 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.

We ran *t*-tests on Accuracy for the test phases and the follow up. All analyses were evaluated at the $\alpha = .05$ level of significance. As previously done the Bayes factor was reported for each analysis. The data set is available via the following link: https://osf.io/xjdt4/.

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 ($M = 933$ ms, $SD = 164$) than for low-contingency trials ($M = 963$ 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, $M_{low\text{-}high\ contingency} = 43.2$, $SD = 111$, compared to the incidental group, $M_{low\text{-}high\ contingency} = 18.5$, $SD = 81.3$.

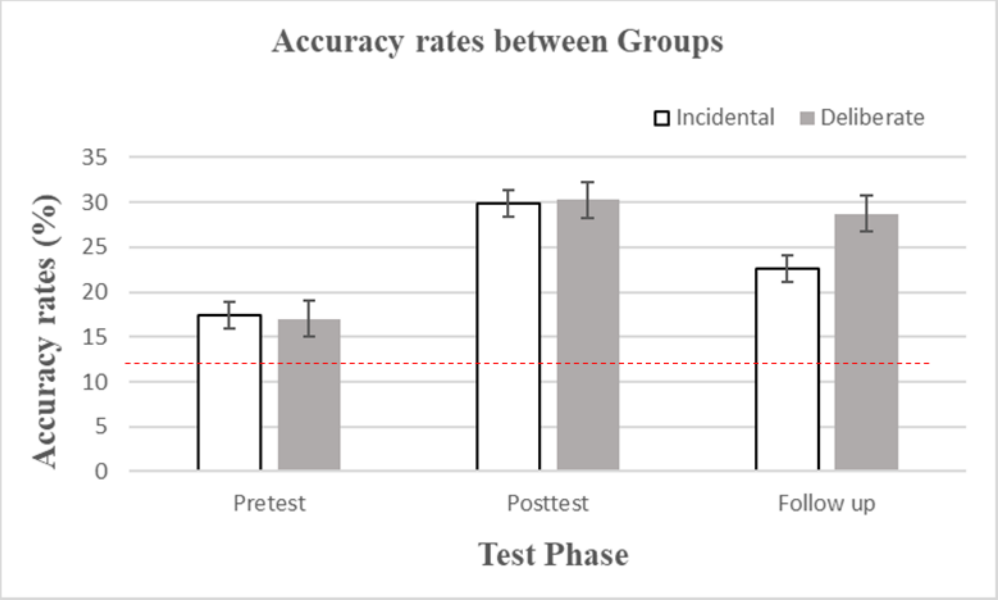
Error Rates

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, $M_{high\text{-}trials} = 12.8\%$, $SD = 7.47$, $M_{low\text{-}trials} = 15.0\%$, $SD = 9.53$; deliberate group, $M_{high\text{-}trials} = 13.5\%$, $SD = 9.59$, $M_{low\text{-}trials} = 16.0\%$, $SD = 12.3$).

Test phases

For the subjective awareness question, 59.25% (80 of 135) of the participants noticed the contingencies between tone sand 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, $t(133) = -.154$, $p = -.462$ $d = -.026$, $BF_{10} = .187$, however the deliberate group performed significantly better than the incidental group in the follow up $t(133) = 1.981$, $p = .0496$, $d = .342$, $BF_{10} = 1.091$.

Figure 21 - 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.

Accuracy rates were significantly higher in posttest compared to pretest in both groups, as shown in Figure 21: 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 $t(61) = 4.84$, $p < 001$, $d = 615$, $BF_{10} > 100$.

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 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 experiments investigated the role of incidental learning in identifying pitches. These results support the idea that incidental learning can help in the acquisition of pitch-label association and pitch identification. Indeed, very robust 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. Although all the nonmusicians were able to acquire pitch-label associations at some level, a big difference was found in the pre- and posttests between the groups. Specifically, the multiple-cues group, which was exposed to the combination of notes 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 and 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 only tone-cue group. Furthermore, these results seem to be inconsistent with the idea that using auditory and visual information can help learning even further. For instance, some previous studies (see Mishra, 2014 for a review) suggested the idea that auditory information have a role in sight-reading performance.

Although, certainly both auditory and visual components are important to learn how to sight-reading, the combination of the two might not be the best option when it comes to the acquisition of new information. These results support the idea of the presence of an overshadowing effect, therefore when presenting two predictive cues the association between one of the two presented stimuli and the target decrease, thus badly influencing the acquisition of the association.

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, the outcomes demonstrated that musicians are also able to strength 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. As an addition, in Experiment 2 we controlled for possible influence of the SMARC effect (Rusconi et al., 2006) on the pitch-label acquisition. Overall musicians seemed to learn better the pitch-label association in the compatible then in the incompatible group, suggesting that the spatial compatibility between tones and responses positively influence the learning process. However, surprisingly, the incompatible group reported faster RT than the compatible group in the learning phase. Although at first this may appear to be inconsistent with the facilitation effect of a spatial compatibility between tones and responses, it is worth noting that the incompatible group was asked to answer with a key-label starting from “do” to “si”, that is they were exposed to a more “classical” order of the note names, the one that likely everyone learn at the first grade of elementary school. It is then possible that this key-label position “facilitated” the answer for the incompatible group, resulting in faster RT.

As another addition, in Experiment 2, we also included 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 the 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 more than a stable “pitch template”.

In Experiment 2, we also controlled for the SMARC effect (i.e., spatially coded response for high pitch on the right and low pitch on the left, see Ariga & Saito, 2019; Rusconi et al., 2006). Evidence for learning was observed overall, but spatial compatibility did influence test phase performance. 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. The compatible group responded above chance in both the pre- and posttest and almost doubled their accuracy rates after the learning phase, whereas the incompatible group’s performance was below chance guessing and did not significantly improve between tests. Accordingly, if humans spatially code the pitches (Ariga & Saito, 2019; Rusconi et al., 2006), it may not seem surprising that the incompatible group was not able to properly learn pitch-label association. It is possible, indeed, that the incongruency between pitches and the spatial location 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 in the compatible group), but spatial incompatibility seems to make this learning more difficult.

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 do 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., 2022b; Schmidt & De Houwer, 2012d), 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.

As one limitation in our study, we did not use semitones (i.e., the smallest distance between notes in the Western scale), and we only presented whole tones (“do, “ré”, “mi”, “fa”, “sol”, “la”, and “si”). Although this is not in line with previous research, specifically the research that investigated AP acquisition (Van Hedger et al., 2019; Wong, Lui, et al., 2020; Wong, Ngan, et al., 2020), we decided to not use semitones because, contrarily to the previous studies, we recruited nonmusicians and focused on short- and medium-term improvements in pitch detection. However, future research may implement in an incidental learning procedure the use of the semitones. 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, that was done simply because it was not our aim to claim that our procedure can teach AP to nonmusicians. Therefore, although we use a pre/post measures to determine at which degree our participants were able to name the tones before and after the training, we use these scores to measure participants' improvements and not their AP abilities. However, future research more focused on AP acquisition may implement AP tests to investigate at what extents our procedure may benefit AP acquisition.

It is also important to reiterate that the goals of the present work diverge from those of past work on pitch detection learning and, more particularly, learning of AP. As mentioned in the Introduction, much work on this topic has focused on the determinants of AP, with both genetic factors and early music learning being indicated as key factors. Some debate has raged about whether AP (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 AP. Our work asked a notably different question: whether the same sort of rapid learning and automatization observed in (nonmusical) incidental learning procedures can also be observed in a pitch learning context. That is, with a drastically shorter learning procedure, can evidence of robust improvements already be observed both 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 improve and RT and error rate indices are impacted by the acquired contingencies. In other words, participants not only improve their explicit tone naming scores, but this retrieval is fast an 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 experienced in a small amount of time. 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 AP, and whether our approach is more effective than other alternatives.

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.

Chapter 4 - learning and time perception

Learning music through time perception

Claudia Iorio and James Schmidt

in

collaboration with Sonja Kotz

Introduction

Recently, there is an increasing interest in temporal perception and its influence on learning processes. That is, attention is focused on the way in which presentation rates (i.e., the time between the presentation of one stimulus and the next during a task) may influence the way in which we perceive and encode information, thus influencing the learning process itself (Destrebecqz & Cleeremans, 2003; Heun et al., 1998; Kiss et al., 2019; Kóbor et al., 2021; Staresina & Davachi, 2009; Wlotko & Federmeier, 2015). For instance, in an artificial grammar task, Selchenkova et al. (2014) asked participants to judge whether a pitch was in tune or out of tune in a sequence of pitches. The pitches were presented either in a temporally regular manner (i.e., stimulus presentation followed a metrical temporal pattern, meaning that each pitch onset was easily predictable) or in an irregular manner, in which the interval between stimuli was randomly varied on a trial-to-trial basis, making them less predictable. Participants were able to learn the pitch grammar better in the regular presentation condition than when the stimuli followed an irregular presentation. The authors suggested that external regularities can help listeners to develop perceptual expectations about the temporal occurrence of future tones, thus facilitating the learning of the pitch grammar (see also Geiser et al., 2012; Lange, 2009; Schmidt-Kassow et al., 2009; Schwartze et al., 2011).

However, most of the research presented above investigated this topic in sequence learning procedures (e.g., when participants are asked to identify a deviant tone in a sequence of tones). Of particular interest for this report is the role of temporal structure and predictability in paired associate type learning, specifically in a contingency learning procedure, in which we test for the basic human ability to learn the relationship between two or more co-occurring events in the environment (Schmidt, 2012; Shanks, 2007, for a review).

Contingency learning and temporal perception

Contingency learning is one way to incidentally (i.e., without the intention to learn) learn the relationships between events presented 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). For instance, in the color-word contingency learning procedure (Schmidt et al., 2007) participants are exposed to regularities between a cue (e.g., word) and a target stimulus (e.g., color). For example, the word “move” might be presented frequently in blue (high-contingency trials), but rarely in green or red (low-contingency trials). Although participants are not instructed to learn the associations between the stimuli, but because they are mostly exposed to high-contingency trials, they do learn the contingencies. Interesting, this learning is extremely rapid: after few trials, responding is robustly faster and more accurate to high-contingency trials than to low-contingency trials (Schmidt et al., 2007; for related learning procedures, Carlson & Flowers, 1996; Miller, 1987; Mordkoff & Halterman, 2008; Musen & Squire, 1993), which we refer to as the contingency effect.

In contingency learning research, the role of time perception has been investigated in a few reports, especially with regards to the temporal contiguity hypothesis (Buehner, 2005). According to this perspective, temporal perception can influence the way we bind together the items, that is, the acquisition of the contingencies is due to the closeness in time between the cue and the target. For example, it is easier to notice that two stimuli tend to be presented one after another if the second stimulus is presented very shortly after the first. However, Schmidt and De Houwer (2012) proposed that the contingency learning effect (low minus high contingency trials) is not directly influenced by temporal contiguity. When words were presented before the target color, increasing the interval of time between the word and color presentation did not seem to meaningfully impact the size of the learning effect. Their results support a temporal insensitivity hypothesis, suggesting that for a performance task (when an explicit judgment is not required) temporal contiguity does not seem to notably influence the size of the contingency effect.

Surely the temporal insensitivity hypothesis suggests that contingency learning is insensitive to temporal manipulations, however it can be argued that a different temporal manipulation (e.g., predictability of the closeness in time between the stimuli) may interfere with the acquisition process and thus influence the size of the contingency effect. Accordingly, the temporal coding hypothesis suggests that temporal information (the “when”) is encoded at the same time as the identity of the events (the “what”; Greville & Buehner, 2010; Matzel et al., 1988), and this may actually have an effect on the acquisition process itself. Furthermore, Balsam et al. (2010) suggested that learning involves extracting the temporal structure of the events and that this temporal information may influence learning of the contingencies itself.

Here it is proposed to investigate the role of time in associative learning beyond the idea of temporal contiguity. In particular, it is argued that predictability in stimuli presentation may help to enhance the learning process in contingency learning. According to the Dynamic Attending Theory (DAT; Jones & Boltz, 1989) rhythmic patterns entrain the attentional system to focus on the regularities and predict the stimuli. Therefore, in an “unpredictable” situation the attentional system should be disrupted and the learning process itself should be less efficient. The aim of this research is to investigate the role of temporal predictability in contingency learning in the framework of the DAT. It is hypothesized that a temporally predictable presentation of the stimuli should facilitate the learning process itself. Therefore, a larger contingency effect is expected in the regular metrical presentation condition compared to the irregular condition.

Music contingency learning procedure

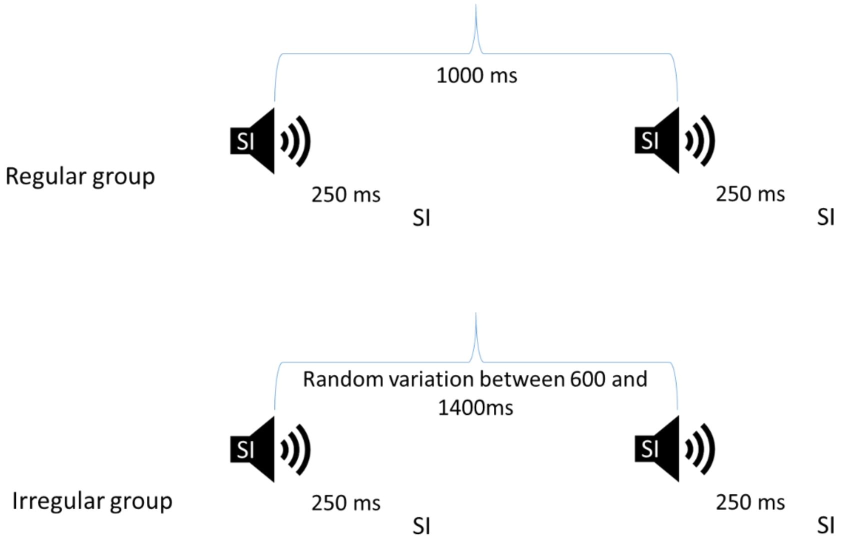
The music contingency learning procedure (Iorio et al. 2021) has been recently developed to study the acquisition of new complex material in an incidental way (i.e., without the goal to intentionally learn new information) in a musical context. This procedure, inspired by the color-word contingency learning paradigm of Schmidt et al. (2007; for reviews, see MacLeod, 2019; Schmidt, 2021) and the music Stroop task of Grégoire et al. (2013, 2014b, 2014a, 2015, 2019), applies knowledge from human contingency learning research (typically with very arbitrary stimuli) to the more practical case of music learning. In the music contingency learning procedure, participants are asked to respond to a target (i.e., note name) while ignoring a predictive cue. In Iorio and colleagues (2022), the predictive cue was a nontarget note position on a musical staff (related to learning to sight-read music notation), but in the current work the predictive cue is a tone (related to learning to detect pitches by ear). Learning can occur, because each note name (e.g., “do”) is presented more often with one tone (e.g., the tone “do”; high-contingency trial, presented 90% of the time) than the others tones (e.g., the tone “mi”; low-contingency trials, presented 10% of the time, see Figure 24 for an example of high and low trials), is illustrated in Table 4. The goal of the procedure is to see if participants can learn to identify pitches by ear, something which most people (including most musicians) are thoroughly unable to do. Some people already do possess this ability, which is referred to as Absolute Pitch (AP; defined as the ability to automatically and effortless identify and name pitches without any external reference; see Deutsch, 2013; Levitin, 1994; Levitin & Rogers, 2005; Takeuchi & Hulse, 1993). Our task is used to assess whether those who do not possess AP are able to learn to identify pitches by ear.

Table 4 - Relative presentations of tones with the note names.

Note Names	*Tones*			
	do	**ré**	**mi**	**fa**
do	84	1	1	1
ré	1	84	1	1
mi	1	1	84	1
fa	1	1	1	84

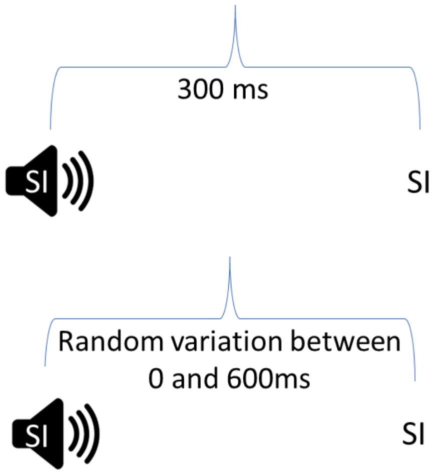
After a very short time, although participants are often not aware of the contingencies, learning of the regularities between tones and note names occurs, allowing participants to anticipate the likely response on the basis of the predictive cue, resulting in a contingency learning effect (i.e., faster RTs in high- than low- contingency trials). However, in the music contingency procedure above, the regularities presented refer to the co-occurrence of the pairings (i.e., high-contingency trials are presented much more often than low-contingency trials). Here, our aim is slightly different. We want to investigate whether the regularities in the temporal structure between stimulus presentations can influence the learning process and therefore the size of the above-described learning effect. As such, we introduce two different temporal manipulations.

Figure 22 - Difference in the temporal structure between the regular and the irregular groups in Experiment 1.



In the first study, we manipulated the intertrial interval (ITI; i.e., the interval between trials) to be either temporally regular or temporally irregular, as illustrated in Figure 22. The tone (predictive cue) is played, then after 250 ms the note name (target) is shown. In the regular group, participants have 1000 ms to respond to the target. On the contrary, the irregular group is characterized by a disruption in the regularities of the time structure. That is, participants have a variable amount of time (randomly selected between 600 and 1400 ms) to respond before the start of the next trial. This irregularity might preclude the possibility to extract temporal information and thus determine a temporal predictability, resulting in worse performance compared to a situation in which the trials follow a regular temporal structure. It is expected that the size of the contingency learning effect will be bigger in the regular condition compared to the irregular one. Specifically, participants in the regular condition are expected to show a contingency effect of comparable size to the one shown in previous studies (Iorio et al., 2022, for the music contingency learning procedure; Schmidt et al., 2007, for the color-word contingency learning procedure), whereas participants in the irregular condition are expected to show a reduced learning effect. This would suggest that there is a beneficial effect of a regular temporal structure over an irregular one in entraining the attentional system and thus benefit the acquisition process.

Figure 23 - Difference in the temporal structure between the regular and the irregular groups in Experiment 2.



However, contrarily to previous research (Destrebecqz & Cleeremans, 2003; Heun et al., 1998; Kiss et al., 2019; Kóbor et al., 2021; Selchenkova et al., 2014; Staresina & Davachi, 2009; Wlotko & Federmeier, 2015) that mostly focused on sequential learning, contingency learning is a kind of associative learning. That is, it is possible that, in contingency learning, attention is directed toward binding together the stimuli within the trial more than between the trials. Therefore, in the second experiment, we manipulated the temporal presentation between the cue and the target (Stimulus Onset Asynchrony, or SOA). We compared again a temporally regular stimulus presentation (i.e., the cue appeared always 300 ms before the target) to a temporally irregular presentation (i.e., the time between the cue and the target was randomly selected between 0 and 600 ms), as illustrated in Figure 23. It was expected that a bigger contingency learning would be observed in the regular group compared to the irregular one. However, it is possible that, consistent with the insensitivity hypothesis (Schmidt & De Houwer, 2012b), the temporal manipulation would not influence the acquisition of the contingencies. This finding would contrast with the previous results in (admittedly different) procedures (Geiser et al., 2012; Schmidt-Kassow et al., 2009; Selchenkova et al., 2014) and the DAT (Jones & Boltz, 1989), which would both predict that a temporally regular presentation between cue and target will benefit the acquisition of the contingencies more than a temporally irregular presentation.

Experiment 1

The aim of the current experiment was to investigate whether manipulating the ITI would influence the size of the contingency effect. For this purpose, we used a modified version of the musical contingency learning mentioned above (Iorio et al., 2022; Schmidt et al., 2022). Specifically, we modified the temporal structure of the trials to create temporally regular trials vs. temporally irregular trials.

Figure 24 - Examples of high/congruent and low/incongruent trials.

Congruent/high trial

Incongruent/high trial

Note. In the left panel is an example of congruent/high contingency trial. It defined as a congruent trial because the tone matches the note name, and also high-contingency because this association was presented much more often during the task. In the right panel is an example of incongruent/low trial, in which the tone does not match the note name and this pairing is presented infrequently during the task.

We then compared performance in the music contingency task between a regular and an irregular group (Figure 23). Based on the DAT (Jones & Boltz, 1989), the attentional system is guided by internal oscillations entrained by a temporally regular presentation and this should induce temporal expectation of future events. In other words, because the stimuli are presented at a constant time, the attentional system is guided to pay attention to these regularities, which makes it easier to learn them and thus to predict future events. Our primary hypothesis is that both groups of nonmusicians will incidentally learn the pitch-label associations. In particular, we anticipate faster (and possibly more accurate) responses to congruent (high-contingency) trials than to incongruent (low-contingency) trials (see Figure 24 for an example of high- and low- contingency trials). However, we also anticipate an effect of the group factor. We expect larger learning effects in the regular group compared to the irregular group (i.e., an interaction between contingency and group).

Method

Participants

Participants (107) were recruited online on Prolific.co (they received monetary compensation of about 4£) and they were randomly assigned to one of the two groups described above (53 participants in the regular group and 54 participants in the irregular group). Inclusion criteria were: i) being able to understand English, ii) being between 18-30 years old, and iii) not being a musician or a singer. As previously done (Iorio et al., 2022) and based on the literature that suggests that participants reporting a score higher than 60% in a pitch identification task may be AP possessors (Levitin & Rogers, 2005; Miyazaki, 1988), we took some measure about AP. Following these criteria, we excluded 11 participants that had accuracy rates above 60% in the pretest (highest performance 83.33%), leaving 48 participants in the regular group and 48 in the irregular group.

All participants signed 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

We used Psytoolkit, a web-based software that allows reliable RT recording, as shown from previous research (Stoet, 2010, 2016), also with music stimuli (Armitage & Eerola, 2020), to code and to run the experiment. The auditory stimuli were created using Audacity software with the lowest tone being the “do” (or “C”) note at the frequency of 261.63 Hz and the highest tone being the “fa” (or “F”) note at the frequency of 349.29 Hz. We used the first 4 tones from the C Major musical scale (i.e., do, ré, mi, fa). We note that this is slightly different than our previous studies (Iorio et al., 2022), which used all 7 tones from one octave of the same C Major scale. This change was made due to the temporal constraints of the present version of the task. Response times to a seven-choice task would be much too slow, on average, whereas a four-choice task seemed more reasonable.

During the main parts of the experiment, participants responded with the T-Y-U-I keys, that were labelled according to the sequence of the musical scale on the music staff (i.e., do, ré, mi, and fa, respectively, using the French note names). The “O” and “N” keys were additionally used to answer “Oui” (Yes) or “Non” (No) to the subjective awareness question (discussed below), and the spacebar was used to begin each phase from the instruction screens.

Participants were randomly assigned to one of the two groups. In the learning phase, both groups were asked to respond to the note names (target). However, while in the regular group the ITI (intertrial-interval) was always identical (i.e., 1000 ms), in the irregular group the ITI randomly varied from trial to trial (i.e., random variation between 600 ms and 1400 ms). The procedure was otherwise identical for the two groups.

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

> *“Do you have absolute 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 (40 trials in each practice phase) in which participants practiced and automatized the note name-to-key assignments. 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 (*do*, *ré*, *mi*, or *fa*) was then presented in the center of the screen until response (no time limit). An on-screen key reminder (T, Y, U, I) was added throughout the first practice phase to help participants to learn the note name-to-key assignments. If participants answered correctly, the next trial began immediately, otherwise, 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.

Following this, participants were presented with a tone naming task, in which they had to guess the name of the tone (no limit time; 12 trials). This pretest phase measures the ability of the participant to discriminate (e.g., better-than-chance guessing) between experienced and unexperienced events (Cheesman & Merikle, 1984), and it allowed us to assess the ability of participants to identify tones prior to learning. The same test phase was then added at the end of the experiment (i.e., after the learning phase, to be described next) as a posttest. As previously mentioned, the current experiment was conducted with nonmusicians, who should normally have no pitch detection skills in the absence of music training, but the test phases, both before and after the learning phase, allowed us to both (a) screen for undisclosed preexisting knowledge and for AP possessors and (b) to establish a control for pre/post improvement scores.

Immediately after the pretest phase, participants started the learning phase that differed between the groups only with regard to the temporal structure of the stimulus presentation (as shown in Figure 22). On each trial, a tone played for 250 ms. The note name was then written in the center of the screen. Based on the group they were in; participants had either 1000 ms or a random interval between 600 and 1400 ms to respond. If participants responded incorrectly or failed to respond, the note name was replaced with “XXX” in red for 500 ms before the beginning of the next trial. In total, there were 420 trials in the learning phase, randomly ordered (without replacement), and a contingency manipulation of 90% (Schmidt et al., 2022) congruent pairings (e.g., the tone “do” with the note name “do”; high-contingency trials) and 10% incongruent trials (e.g., the tone “do” with the note name “mi”; low-contingency trials; see Table 4 for the proportion of the contingency manipulation). The congruency (or contingency learning) effect was measured as the difference between low- and high- contingency trials.

Following the main learning phase, we asked participants whether they noticed the regularities during the learning phase as a measure for contingency awareness. In particular, 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:

> *“During the third part of this experiment, note names were presented with a tone. 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. The instructions for these phases were:

> *“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.”*

Data Analysis

Only data from the learning and the test phases were used for the analysis. Specifically, correct RTs and errors rates from the learning phase were used to conduct a repeated measures ANOVA to assess the overall main effects of Contingency, Group, and the interaction between them. We eliminated trials in which participants failed to respond in time (based on the temporal structure, the deadline was either 1000 ms or a random interval between 600 and 1400 ms). Accuracy rates from the test phases were analyzed to assess whether participants responded above chance (the chance guessing rate was 1/4 or 25%) and whether there was a significant improvement in identifying pitches pre and post learning phase. All analyses were evaluated at the α = .05 level of significance. Additionally, we consistently report 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 reported the Bayes factor BF_{10}, with values between 3 and 10 supporting a moderately strong evidence for the alternative hypothesis (H1; Doorn et al., 2021). The data set is available via the following link: https://osf.io/msx52/.

Results

Response times

The repeated measure ANOVA on RTs with Contingency (high vs. low) and Group (regular vs. irregular) as factors showed a significant main effect of Contingency, $F(1,94) = 85.42$, $p < .001$, $\eta^2 = .476$, $BF_{10} > 100$, indicating faster responses for high-contingency trials (M = 550 ms, SD = 46.7) than for low-contingency trials (M = 572 ms, SD = 47.9). However, the main effect of Group, $F(1,94) = 2.93$, $p = .090$, $\eta^2 = .030$, $BF_{10} = .506$, and the interaction between Contingency and Group, $F(1,94) = 1.03$, $p = .312$, $\eta^2 = .011$, $BF_{10} = .112$, were not significant. The contingency effect (i.e., the difference between high- and low- contingency trials) was significant for the regular group, $t(47) = 5.73$, $p < .001$, $d = .827$, $BF_{10} > 100$ (M_{high}=560 ms, SD = 42.0, M_{low}=579 ms, SD = 43.7), and the irregular group, $t(47) = 7.36$, $p < .001$, $d = 1.06$, $BF_{10} > 100$ (M_{high}=541ms, SD = 49.7, M_{low}=566 ms, SD = 51.2).

Error rates

The repeated measure ANOVA on error rates with Contingency (high vs. low) and Group (regular vs. irregular) as factors showed a significant main effect of Contingency,

$F(1,94) = 33.44$, $p < .001$, $\eta^2 = .262$, $BF_{10} > 100$, indicating higher error rates for low-contingency trials ($M = 16.7\%$, $SD = 9.89$) than for high-contingency trials ($M = 13.7\%$, $SD = 8.58$). Both the main effect of Group, $F(1,94) = 6.79$, $p = .011$, $\eta^2 = .067$, $BF_{10} = 9.40$, and the interaction between Contingency and Group, $F(1,94) = 5.51$, $p = .021$, $\eta^2 = .055$, $BF_{10} = 7.99$, were significant (regular group's mean, $M_{high} = 12.0\%$, $SD = 7.48$, $M_{low} = 13.8\%$, $SD = 7.40$; irregular group's mean, $M_{high} = 15.4\%$, $SD = 9.34$, $M_{low} = 19.6\%$, $SD = 11.2$).

Test phases

We conducted analysis on the accuracy rates for the pre and the posttests. *T*-tests showed no significant differences in accuracy rates between the groups for the pretest, $t(94) = -1.68$, $p = .0.96$, $d = -.343$, $BF_{10} = .743$ ($M_{regular\text{-}group} = 27.6\%$, $SD = 16.8$; $M_{irregular\text{-}group} = 33.2\%$, $SD = 15.6$), but the difference in accuracy was significant in the posttest, $t(94) = -2.20$, $p = .030$, $d = -.450$, $BF_{10} = 1.79$ ($M_{regular\text{-}group} = 37.8\%$, $SD = 20.2$; $M_{irregular\text{-}group} = 48.3\%$, $SD = 25.8$). The regular group performed not significantly above chance guessing (i.e., 25%) in the pretest $t(47) = 1.08$, $p = .287$, $d = .155$, $BF_{10} = .270$, $M = 27.61\%$, $SD = 16.8$, but in the posttest, $t(47) = 4.41$, $p < .001$, $d = .636$, $BF_{10} > 100$, $M = 37.8\%$, $SD = 20.2$ performance was significantly above chance guessing. The irregular group reported performance significantly above chance guessing in both the pretest, $t(47) = 3.62$, $p < .001$, $d = .523$, $BF_{10} = 38.9$, $M = 33.2\%$, $SD = 15.6$, and the posttest, $t(47) = 6.25$, $p < .001$, $d = .902$, $BF_{10} > 100$, $M = 48.3\%$, $SD = 25.8$. Both the regular, $t(47) = 3.36$, $p = .002$, $d = .485$, $BF_{10} = 19.3$, and the irregular, $t(47) = 3.39$, $p = .001$, $d = .490$, $BF_{10} = 21.3$, groups showed significantly higher accuracy rates in the posttest compared to the pretest. We also ran a repeated measures ANOVA with Test (pre and post) and Group (regular vs. irregular) as factors. While the main effect of Test, $F(1,94) = 22.06$, $p < .001$, $\eta^2 = .190$, $BF_{10} > 100$, and Group, $F(1,94) = 6.81$, $p = .011$, $\eta^2 = .068$, $BF_{10} = 2.49$, were significant, the interaction between Test and Group was not significant, $F(1,94) = .812$, $p = .370$, $\eta^2 = .009$, $BF_{10} = .313$.

Discussion

Experiment 1 aimed to investigate the effect of temporal structure on the learning process. According to the DAT (Jones & Boltz, 1989), it was hypothesized that temporally regular trials may help guide the attentional system to perceive and encode the stimuli better than temporally irregular trials. Specifically, in the context of contingency learning, it was hypothesized that nonmusicians would have been able to learn pitch-label associations, with a contingency learning effect that would be bigger for the regular group compared to the irregular group. Our results showed that, in line with the first hypothesis, all the nonmusicians were able to learn pitch-label associations and the contingency effect was significant for both the regular and the irregular groups. Also, there was a general tendency to make more errors for low-contingency trials than for high-contingency trials, and this was also significantly different between the groups, with the irregular group making more errors than the regular one.

Furthermore, participants' performance in the tests in Experiment 1 significantly improved between the pre- and posttest, although there was no significant difference between the groups, inconsistent with the original hypothesis.

Experiment 2

Overall, the main goal of Experiment 2 was similar to Experiment 1. That is, we wanted to study whether manipulation in the temporal structure could influence the size of the contingency effect. However, contrarily to Experiment 1, here we manipulated the temporal intervals between the cue and the target (SOA). We expected to find a significant contingency effect for both the regular rand the irregular groups, suggesting that all the participants are able to learn pitch-label associations. However, the interaction between groups is less clear. While a nonsignificant interaction may suggest that temporal manipulations do not affect contingency learning, a significant interaction would provide evidence for the influence of time in associative learning. Our hypothesis was that the learning effect would be larger with a temporally regular presentation.

Method

Participants

Using Prolific.co, we recruited 98 participants randomly assigned to one of the two groups described above (41 participants in the regular group and 57 participants in the irregular group). Participants that met the inclusion criteria (being able to understand English, being between 18-30 years old, and not being a musician or a singer) received a monetary compensation (4£) for their participation. Again, following the same criteria used in the first experiment mentioned above, 13 participants that had accuracy rates above 60% in the pretest (highest performance 70.69%) were excluded, leaving 35 participants in the regular group and 50 in the irregular group.

All participants signed 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 same exact structure from Experiment 1 was implemented in Experiment 2, except for one modification. In the learning phase, in Experiment 2 we manipulated the interval between the cue and the target (SOA) instead of manipulating the ITI as it was in Experiment 1. Therefore, while in the regular group the cue stayed on the screen for 300 ms before the target appeared, in the irregular group the SOA between the cue and the target was randomly selected between 0 and 600 ms.

Data Analysis

The data analysis followed the same criteria used in Experiment 1. The data set is available via the following link: https://osf.io/msx52/.

Results

Response times

We ran a repeated measure ANOVA on RTs with Group (regular vs. irregular) and Contingency (high vs. low) as factors that showed a significant main effect of Contingency, $F(1,83) = 72.41$, $p < .001$, $\eta^2 = .466$, $BF_{10} > 100$, indicating faster responses for high-contingency trials (M = 639 ms, SD = 100) than for low-contingency trials (M = 670 ms, SD = 101). However, the main effect of Group, $F(1,83) = .860$, $p = .356$, $\eta^2 = .010$, $BF_{10} = .649$, and the interaction between Contingency and Group, $F(1,83) = 1.45$, $p = .232$, $\eta^2 = .017$, $BF_{10} = .417$, were not significant. The contingency effect (i.e., the difference between high- and low-contingency trials) was significant for the regular group, $t(33) = 6.47$, $p < .001$, $d = 1.11$, $BF_{10} > 100$ (M_{high}=624 ms, SD = 105, M_{low}=660 ms, SD = 110), and the irregular group, $t(50) = 5.66$, $p < .001$, $d = .793$, $BF_{10} > 100$ (M_{high}=649 ms, SD = 96.3, M_{low}=676 ms, SD = 95.7).

Error rates

We ran a repeated measure ANOVA on error rates with Contingency (high vs. low) and Group (regular vs. irregular) as factors that showed a significant main effect of Contingency, $F(1,83) = 24.7$, $p < .001$, $\eta^2 = .229$, $BF_{10} > 100$, indicating higher error rates for low-contingency trials (M = 10.6%, SD = 9.28) than for high-contingency trials (M = 8.60%, SD = 7.87). However, both the main effect of Group, $F(1,83) = 1.18$, $p = .281$, $\eta^2 = .014$, $BF_{10} = .666$, and the interaction between Contingency and Group, $F(1,83) = .0001$, $p = .991$, $\eta^2 = .000$, $BF_{10} = .187$, were not significant (regular group's mean, M_{high} = 7.40%, SD = 3.88,

M_{low}=9.42%, *SD* = 4.81*;* irregular group's mean, M_{high} = 9.41%, *SD* = 9.61, M_{low} = 11.4%, *SD* = 11.3*).*

Test phases

We conducted analyses on the accuracy rates for the pre- and the posttests. *T*-tests showed no significant differences in accuracy rates between the groups for the pretest, *t*(83) = .041, *p* = .967 *d* = .009, BF_{10} = .231, and in the posttest, *t*(83) = .336, *p* = .737, *d* = .074, BF_{10} = 242. Both the regular and irregular groups showed performance significantly above chance guessing (i.e., 25%) in the posttest: regular group, *t*(33) = 4.83, *p* < .001, *d* = .828, BF_{10} > 100, *M* = 45.1%, *SD* = 24.3*;* irregular group, *t*(50) = 5.44, *p* < .001, *d* = .761, BF_{10} > 100, *M* = 43.3%, *SD* = 24.0, but the performance were not significantly above chance guessing in the pretest: regular group, *t*(33) = 1.27, *p* = .214, *d* = .218, BF_{10} = .383, *M* = 28.9%, *SD* = 18.0*;* irregular group, *t*(50) = 1.53, *p* = .133, *d* = .214, BF_{10} = .451, *M* = 28.8%, *SD* = 17.6 ,. Both the regular, *t*(33) = 3.59, *p* = .001, *d* = .616, BF_{10} = 30.5, and the irregular, *t*(50) = 4.16, *p* < .001, *d* = .582, BF_{10} > 100, groups showed significantly higher accuracy rates in the posttest compared to the pretest.

As in Experiment 1, we ran a repeated measures ANOVA with Test (pre and post) and Groups (regular vs. irregular) as factors. While the main effect of Test, *F*(1,83) = 29.50, *p* < .001, η^2 = .263, BF_{10} > 100, was significant, the main effect of Group, *F*(1,83) = .068, *p* = .794, η^2 = .001, BF_{10} = .208, and the interaction between Test and Group, *F*(1, 83) =.083, *p* = .7773, η^2 = .001, BF_{10} = .222, were not significant.

Discussion

In Experiment 2, we changed the type of temporal manipulation to study whether a predictable interval between the cue a target compared to an unpredictable interval could benefit the acquisition of the contingencies. Although all the nonmusicians were able to learn pitch-label associations as indicated by the posttests scores, we did not find a significant interaction between contingency and group. These results together with the data from Experiment 1 seem to suggest no influence of temporal manipulations on contingency learning.

General discussion

The aim of these studies was to investigate the role of temporal perception in human contingency learning. According to the DAT (Jones & Boltz, 1989), temporally regular patterns help the attentional system to focus on the stimuli and to encode them better. Based on this assumption we compared performance in a musical contingency learning procedure between a regular and an irregular group. In Experiment 1, although our results showed that all the nonmusicians were able to learn the pitch-label associations presented in the task, contrarily to the previous research mentioned above we did not find a significant difference between the groups, suggesting that in associative learning the temporal structure may not influence the way we learn the contingencies.

It is worth noting, however, that there are a few small differences between our procedure and the tasks previously mentioned. As a first difference, in the current research the task is based on an associative learning paradigm, rather than a sequence learning task. It may be possible that we did not find the same results as before because in associative learning people are more focused on the timing between the cue and the target than the general temporal structure of the trials. In other words, in the previous research people were dealing with one level of predictability (that is the time between the stimulus presentations), but in the current research participants are exposed to two levels of predictabilities, the first being the contingencies between stimuli, the second being the temporal structure of the task (regular vs. irregular). It is possible that in this kind of associative learning task people are more focused on binding together the stimuli and do not perceive the timing between the trials, therefore they are not influenced by the changes in the temporal structure of the trials. To test this hypothesis, in Experiment 2 we manipulated the interval between the cue and the target. However, although all the participants learned the pitch-label associations, we did not find a significant difference between the groups.

Overall the collective results from Experiments 1 and 2 seem to be more in line with the temporal insensitivity hypothesis proposed by Schmidt and De Houwer (2012b), that is, it may be the case that in performance task temporal structure does not influence the way we perceive the stimuli or the way in which we bind them together.

As a curiosity, in Experiment 1 we observed a peculiar result in the performance data. That is, participants responded faster (at least numerically) and made more errors in the irregular group compared to the regular one. This was a little surprising considering that in the previous research cited above, participants in the regular condition responded faster than participants in the irregular condition. However, our results seem to be in line with the Speed Accuracy Trade off effect (Donkin et al., 2014), that is, faster responses tend to be less accurate than slower responses and people seem to be capable of choosing to make faster responses even if this makes them more prone to errors. Specifically, it is possible that in our tasks participants in the irregular group felt that they had on average less time to respond (e.g., due to the occasional trials with a very short response deadline) and they therefore decided to try to respond as fast as possible (as they were asked in the instructions) even if this meant that they made more errors.

As one limitation, in the current report we used only behavioral measures. It may be argued that a difference between groups can exists at the brain level. Therefore, future studies may add EEG measures or other neuroscientific approaches to investigate temporal processing mechanisms in the context of contingency learning. To conclude, we consider these experiments as another step to a better understanding of the role of time in associative learning tasks. However, more work is needed in order to clarify the relationship between temporal perception and contingency learning.

Chapter 5 - General discussion

The main goal of this thesis was to apply knowledge from human contingency learning research to the music domain. Specifically, I tested the hypothesis that an incidental learning procedure could benefit the acquisition and automatization of musical subskills useful to learn how to play. The main results showed that both in the visual domain (chapter 2) and in the auditory domain (chapter 3) it is possible to apply contingency learning rules to the music domain. Specifically, nonmusicians were able to easily and rapidly learn the associations between the cues and the target (e.g., between the note name and the note position, or the tone and the note name) and successfully used this information in performance and explicit-judgmental tasks.

Furthermore, in chapter 4 I investigate the role of temporal perception in contingency learning. Specifically, I tested the hypothesis that a metrical regular temporal presentation can benefit the acquisition of the contingencies. Surprisingly, the results did not confirm the original hypothesis, suggesting either that temporal perception does not play a role in contingency learning or that further research is needed to explain the role of temporal perception and its effect in contingency learning.

Sight-reading and automatization

The first set of studies, in chapter 2, investigated whether nonmusicians could incidentally learn musical skills needed for sight-reading, that is the ability of reading and of performing a musical piece without previously having practiced it. Sight-reading is considered to be a difficult task that involves not only the ability to decode music symbols (i.e., those used in music notation) but also to transform these symbols into appropriate motoric actions to be able to reproduce the written score on the instrument (Gudmundsdottir, 2010; Kopiez & In Lee, 2006, 2008; Lehmann & Kopiez, 2009).

This is a quite useful and required skill for musicians, however, it is usually lacking among music students (Hargreaves, 1986; Mills & McPherson, 2006; Scripp, 1995). One reason for this lack of skill may be the way this skill is taught. For the most part, explicit tutoring and deliberate practice are, indeed, used to teach this skill (Ericsson et al., 1993; Ericsson & Harwell, 2019; Hébert & Cuddy, 2006; Lehmann, 1997; Mills & McPherson, 2006; Mishra, 2014). One important difference between musicians with a good level of sight-reading and novices is their ability to automatically read music notation (Crump et al., 2012; Drost et al., 2005; Grégoire et al., 2013; Stewart, 2005; Zakay & Glicksohn, 1985, for musical Stroop procedures).

Although, deliberate practice is useful to achieve expertise (Ericsson et al., 1993; Ericsson & Harwell, 2019; Lehmann, 1997), it does not allow one to experience a large amount of practice in seeing each note and reproducing it. Often, repetition is all that is needed to automatize a skill. In incidental procedures like the one reported above, participants are able to experience many repetitions in a small amount of time (e.g., 336 trials in 15 minutes). In the music contingency learning task implemented in the studies above, on each trial, participants identified a note name written inside of a note on the musical staff. In Experiment 1, each note was presented frequently with the congruent note name (e.g., “do” with the note for “do”, high-contingency trials) and rarely with the incongruent names (e.g., “do” with the note for “fa”, low-contingency trials). The contingency effect for reaction times is computed as the RTs of the low-contingency trials minus the RTs of high-contingency trials. To study the efficiency of incidental learning as opposed to deliberate learning to automatize note name/note position associations I compared performance in the music contingency task between a deliberate learning group, who was instructed to pay attention to the contingencies, and an incidental learning group, who was not aware of the contingencies presented in the task. Although both groups were able to successfully learn the associations between the note name and the positions, deliberate learning seemed to help specifically when it comes to explicit reporting of the knowledge previously acquired. In the note position test following the learning phase, the deliberate learning group reported better performance compared to incidental group. As previously mentioned, this is not surprising when considering that to learn contingencies, being attentive to the predictive dimension is important (Eitam et al., 2009; Jiang & Chun, 2001). However, based on these results it is not possible to conclude whether deliberate training can bring better results compared to a purely incidental training, therefore, more research is needed to clarify the role of intentionality in this kind of learning.

In Experiment 2 and Experiment 3 of chapter 2, I studied some related issues, as for instance in Experiment 2 I wanted to investigate whether previous musical knowledge could influence the size of the contingency effect. Specifically, in this experiment I computed the contingency effect and the congruency effect. The analysis reported both significant contingency and congruency effects. Although the associations were “wrong” in Experiment 2, participants learned them, suggesting a contingency learning effect due to the task and not to previous knowledge. However, the congruency effect found in experiment 2 suggests the presence of some previous musical knowledge at least from some of the participants. Overall, although the results suggested evidence for both the effects, previous knowledge did not seem to negatively influence contingency learning.

In Experiment 3, no contingency effect was measured. Instead, I focused on the role of the SMARC (Spatial–Musical Association of Response Codes) effect, defined as a preference to spatially map lower stimuli on the left and higher stimuli on the right (Ariga & Saito, 2019; Rusconi et al., 2006). The results reported no differences between the congruent, compatible, or the control associations suggesting that in performance tasks the spatial relation between the stimuli does not influence the acquisition of the associations. However, a small difference between the congruent and the control associations was found in the test phase, suggesting that in an explicit-judgement task the spatial position between stimuli can help to influence the responses.

Overall, the evidence gathered from the studies above suggest that incidental learning can help in easily and rapidly acquiring note name/note position associations. As mentioned in the introduction, one reason why this procedure worked is that it allowed for many repetitions of the stimulus-response pairings in a short period of time. This in turn may benefit the automatization of the visuomotor integration useful in sight-reading performance.

Pitch identification and automatization

In chapter 3, the set of studies investigated whether the incidental learning procedure described above could be used to help nonmusicians acquiring pitch-label associations and whether this kind of task could benefit the improvement of pitch identification in musicians. Furthermore, I investigated the role of incidental learning as opposed to deliberate learning (i.e., the process of intentionally trying to acquire new information) for memory consolidation of pitch-label associations. The procedure was almost identical to the one used in chapter 2. Participants were asked to respond to the target (note names) while ignoring the cues (either tones or tones with note positions). Contrarily to chapter 2, however, I used auditory stimuli to investigate the research question in the auditory domain.

In Experiment 1, participants were divided into a multiple-cues group (with both note positions and tones as the cues) and a tone-cue group (only tones as the cues). This was done mainly because previous research has suggested that a combined presentation of both note positions and tones can benefit the acquisition of musical skills such as sight-reading (Mishra, 2014). Thus, we compared two groups that were exposed to different cue-target associations.

Overall, nonmuscians were able to learn pitch-label associations, however, the multiple-cues group showed worse performance, suggesting an overshadowing effect (Kamin, 1969).

As mentioned before, the overshadowing effect refers to the situation in which one outcome follows two stimuli, as for instance the tones and the note position, thus learning of the relation between the outcome and one or both of the stimuli is weaker compared to when only one stimulus precedes the outcome (Kamin, 1969; Pavlov, 1927). Therefore, to consolidate and automatize the tone/note name association, it may be preferable to use one cue instead of two.

In Experiment 2, I studied whether the incidental learning procedure could increase performance in pitch identification with musicians. Again, I obtained a significant contingency effect in the learning phase. In addition, the musicians significantly improved between the pre and posttests after the learning phase. Surprisingly, although previous research suggested that AP development can be related to early musical training (Crozier, 1997; Deutsch et al., 2006; Miyazaki & Ogawa, 2006), in Experiment 2 I did not find significant relations between early music training and pitch identification suggesting that more research is needed to clarify this possible relation.

In Experiment 3, I focused on the role of intentionality in acquiring and remembering pitch-label associations. Although, previous studies suggested that deliberate learning can benefit the acquisition of the contingencies (Schmidt & De Houwer, 2012a, 2012d), in chapter 2, no significant difference between an incidental and a deliberate learning group was found in the learning phase. All the participants were able to learn the contingencies, and as I mentioned in chapter 2, here again I did not find any significant difference between the groups in the learning phase. Furthermore, the deliberate group (instructed to learn the contingencies) showed significantly better performance only in the follow up compared to the incidental group (no instruction to learn the contingencies). Overall, the evidence seems to suggest no substantial differences between incidental and deliberate learning with the exception of the follow up in which deliberate learning seems to benefit more the consolidation of the information acquired compared to incidental acquisition.

Time and learning

In chapter 4, I investigated the relationship between time and learning. Specifically, I wanted to study whether temporal perception can influence the acquisition process. Previous research (Greville & Buehner, 2010; Matzel et al., 1988), indeed, seems to suggest that when we learn new information we encode not only the information itself, but also temporal information, specifically "when" the information occurred. Furthermore, according to the DAT (Jones & Boltz, 1989), metrical temporal intervals may help the attentional system to focus and be prepared to encode new information.

Following these lines of research, I hypothesized that manipulating the temporal presentation of the stimuli could influence the way we learn the contingencies. While a regular presentation may help to entrain the attentional system and to benefit the encoding of the contingencies, an irregular presentation could negatively interfere with the encoding. Specifically, I manipulated the temporal presentation of the stimuli in two different ways: while in Experiment 1 I manipulated the interval rates between the trials, in Experiment 2 interval rates between the cue and the target was changed.

However, contrary to what was hypothesized and to the previous research, both the Experiment 1 and the Experiment 2 results did not show a significant difference between the two temporal groups (regular vs. irregular). Although this is surprising, as previously mentioned, it is worth noting that there are a few differences between these experiments and the previous works (Destrebecqz & Cleeremans, 2003; Heun et al., 1998; Kiss et al., 2019; Kóbor et al., 2021; Staresina & Davachi, 2009; Wlotko & Federmeier, 2015; Selchenkova 2014). For instance, contrary to the previous studies that mainly involved sequential learning, here participants were presented with an associative learning task. That is, contingency learning mostly refers to the action of binding two stimuli together, in order to anticipate the most likely response that will be required. Thus, participants need to create a mental representation of the most likely association (e.g., the tone “do” is more likely to be associated to the note name “do” than to any other note names). Therefore, it is possible to think that attention is focused above all on the cue and the target more than the timing between the trials. However, even manipulating the interval between the cue and the target did not produce any significant differences between the groups.

As an interesting result, in Experiment 1, the irregular group made more errors and thus was less accurate compared to the regular group. Although, I was not explicitly looking for this kind of results, it is interesting to see how temporal information can influence behavior. In this specific case and in line with previous work (Donkin et al., 2014), the participants seemed to choose to respect the instructions they were give (i.e., being as fast and as accurate as possible) and aimed for speed at the cost of accuracy in responding. This may suggest an influence of the time structure at least at some level, that is, in the irregular group, participants may feel that they have too little time to respond. However, overall this do not seem to influence the acquisition of the contingencies.

Limitations and future directions

Overall, the studies reported in the thesis applied well-known law of human behavior from cognitive psychology to investigate whether, from both a theoretical and practical perspective, incidental learning procedures can help nonmusicians to easily and rapidly gain knowledge about musical information useful to learn how to play.

However, although I tried to implement ecological stimuli, for instance, note positions written on the musical staff, the stimuli used as well as the modalities were far from the daily practice a nonmusician is confronted with. For instance, in chapter 2 the focus of the research was to investigate whether it was possible to use an incidental learning procedure to teach nonmusicians how to sight-read. Sight-reading refers to the ability to read and play a new musical score without previous practice. This ability mainly refers to the capacity of visuomotor integration (Kopiez & In Lee, 2006, 2008; Lehmann & Kopiez, 2009, Gudmundsdottir, 2010), that is: I read a music symbol and I am able to reproduce it on the instrument. To be able to study whether incidental learning can help in the automatization of the visuomotor integration, I decided to focus on presenting only seven notes to the participants. Although this was done to simplify the task and help nonmusicians to not feel discouraged when confronted with the experiment, it was far from a real-life situation. Indeed, when a novice musician starts to learn how to read and play music, he/she will be confronted with more than seven notes that can be played in different position on the instrument (depending on the instrument, for instance, the same note “fa” can be played at different heights on the violin). Therefore, as a limitation, our tasks did not consider the variety of stimuli that a novice musician will be confronted with. To be able to better understand the practical use of the incidental learning procedure in helping in improving sight-reading performance, future research can implement a bigger variety of stimuli taking into consideration not only all the twelve tones, but also other rhythmic figures, for instance.

Similar limitations can be applied also for chapter 3. In the pitch identification studies I wanted to investigate the utility of an incidental learning procedure in helping nonmusicians to recognize pitches. As mentioned before, we again used only seven notes instead of the twelve normally used in western music. Although this was done once again to help participants to feel comfortable in carrying out the task, it imposed some limitations on the studies. Therefore, as it was for the visual domain, it is possible to think of implementing a much bigger variety of stimuli to investigate the research question in more ecological settings.

In both the visual and the auditory domains, the use of real-life instruments can be better implemented in future research. For instance, the current studies used a computer keyboard for responding. This is, of course, analogous to a piano keyboard. However, future research might be conducted with instruments that are played in notably different ways (e.g., string or wind instruments).

As another limitation, in the current studies I mostly focused on nonmusicians as the preferred participants for the research. This was done to be able to evaluate the efficacy of the incidental learning procedure with people without previous music knowledge similar to the level of novice music students. However, to better understand the practical utility of this incidental learning procedure, it can be possible to compare different groups with different levels of music expertise as well as to compare this incidental learning procedure to other and more classical tasks to teach sight-reading and pitch identification.

As previously mentioned in chapter 4, I wanted to investigate the role of temporal perception in learning the contingencies. Until now the investigation of time in contingency learning mostly focused on considering whether the closeness in time between the cue and the outcome could influence the perception of proximity and the ability to bind together the stimuli (Buehner, 2005). Here I took into consideration another aspect of time perception. Based specifically on the DAT (Jones & Boltz, 1989), I hypothesized that metrical temporal stimulus presentation could help to synchronize the attentional system and have a positive influence on the acquisition of the contingencies. To do that, I use two types of temporal manipulations: at first I compared a regular ITI to an irregular one, secondly I manipulated the SOA between cue and target. Unlike previous research, however, the results suggested that both type of manipulations do not have an effect on the learning process. In future research it may be interesting to investigate the role of temporal perception in contingency learning in neuroimaging studies to consider whether a possible influence of temporal manipulations on the acquisition of the contingencies is possible at brain level.

Side research projects

During the years of the thesis, I had the opportunity to work with some colleagues on a few side projects. Although these collaborations were not part of the thesis work, they gave me the possibility to work on other aspects of learning closely related to the main topic of my thesis, therefore, even though only briefly, I wanted to present them here.

The studies investigated sequential learning using a very common incidental learning procedure: The Serial Reaction Time Task (SRTT). In this particular task, participants are required to press a button corresponding to the location of a visual stimulus. Unbeknownst to the participant, the positions of the visual stimuli are presented in a repeated sequence. The gradual reduction in RT and errors provides a measure of the learning of the sequence.

In the following experiments the SRTT was implemented to study:

a) The role of musical priming on motor implicit learning
b) The role of Response-Stimulus-Interval (RSI) in learning

Previous studies (Selchenkova et al., 2014) showed that temporally regular auditory presentation induces attention-based temporal expectations that can benefit learning processes, like for instance the implicit learning of an artificial grammar or a language. In our study, we investigated whether rhythmic auditory priming could also benefit learning, in particular, the learning of an implicit motor sequence. To do so, we implemented a musical priming Serial Reaction Time Task (SRTT) in which each block of practice was preceded by a short musical sequence. To study the effect of temporal regularity on the acquisition of the sequence, we compared two groups: while one group was presented with regular musical sequences (with an inter-beat interval of 500 ms), the other one was presented with irregular musical sequences.

We found a significant main effect of learning, but no interaction between regularity (regular vs. irregular) and learning, suggesting that the regularity did not influence the ability of participants to learn the sequence. Although our evidence is in line with classic results from the SRTT literature, the nonsignificant interaction suggests that, in contrast to our hypothesis, regular priming does not seem to have a beneficial effect on motor sequence learning. However, more research is needed to be able to determine the effect of a musical priming on the learning of motor sequences.

In another project, we implemented again the SRTT. However, in this study we aimed to investigate whether different manipulations of RSI could influence sequence learning. To test this, we used the classic structure of the SRTT, similar to what I described above, except that no musical priming was used, and we compared three different conditions: in the first, the RSI was set to 0 ms, in the second the RSI was 250 ms and in the last condition the RSI was set at 750 ms. Moreover, to investigate possible differences between age ranges, we decided to compare performance in the SRTT between children, adults, and the elderly. Data collection is still ongoing for this project

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

Anderson, J., R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*, 369–406.

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://www.frontiersin.org/article/10.3389/fpsyg.2019.02883

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*(4), 321–324. https://doi.org/10.1111/1467-9280.00063

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

Augustinova, M., & Ferrand, L. (2014). Automaticity of word reading: Evidence from the semantic Stroop paradigm. *Current Directions in Psychological Science*, *23*(5), 343–348. https://doi.org/10.1177/0963721414540169

Baharloo, S., Service, S. K., Risch, N., Gitschier, J., & Freimer, N. B. (2000). Familial aggregation of absolute pitch. *The American Journal of Human Genetics*, *67*(3), 755–758. https://doi.org/10.1086/303057

Balsam, P. D., Drew, M. R., & Gallistel, C. R. (2010). Time and associative learning. *Comparative Cognition & Behavior Reviews*, *5*(1), 1–22.

Barry, N. H. (1992). The effects of practice strategies, individual differences in cognitive style, and gender upon technical accuracy and musicality of student instrumental performance. *Psychology of Music*, *20*(2), 112–123. https://doi.org/10.1177/0305735692202002

Barry, N. H. (2007). A qualitative study of applied music lessons and subsequent student practice sessions. *Contributions to Music Education*, *34*, 51–65.

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

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*. 31.

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://www.frontiersin.org/article/10.3389/fpsyg.2012.00294

Brodsky, W., Henik, A., Rubinstein, B.-S., & Zorman, M. (2003). Auditory imagery from musical notation in expert musicians. *Perception & Psychophysics*, *65*(4), 602–612. https://doi.org/10.3758/BF03194586

Brodsky, W., Kessler, Y., Rubinstein, B.-S., Ginsborg, J., & Henik, A. (2008). The mental representation of music notation: Notational audiation. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(2), 427–445. https://doi.org/10.1037/0096-1523.34.2.427

Buehner, M. J. (2005). Contiguity and covariation in human causal inference. *Animal Learning & Behavior*, *33*(2), 230–238. https://doi.org/10.3758/BF03196065

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

Catale, C., & Meulemans, T. (2009). The Real Animal Size Test (RAST): A new measure of inhibitory control for young children. *European Journal of Psychological Assessment*, *25*(2), 83–91. https://doi.org/10.1027/1015-5759.25.2.83

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

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

Crump, M. J. C., Logan, G. D., & Kimbrough, J. (2012). Keeping an eye on guitar skill: Visual representations of guitar chords. *Music Perception*, *30*(1), 37–47. https://doi.org/10.1525/mp.2012.30.1.37

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

Destrebecqz, A., & Cleeremans, A. (2003). Temporal effects in sequence learning. In L. Jiménez (Ed.), *Advances in Consciousness Research* (Vol. 48, pp. 181–213). John Benjamins Publishing Company. https://doi.org/10.1075/aicr.48.11des

Deutsch, D. (2013a). 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. (2013b). *Psychology of Music*. Elsevier.

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

Donkin, C., Little, D. R., & Houpt, J. W. (2014). Assessing the speed-accuracy trade-off effect on the capacity of information processing. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1183–1202. https://doi.org/10.1037/a0035947

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

Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005). When hearing turns into playing: Movement induction by auditory stimuli in pianists. *The Quarterly Journal of Experimental Psychology Section A*, *58*(8), 1376–1389. https://doi.org/10.1080/02724980443000610

Eitam, B., Schul, Y., & Hassin, R. R. (2009). Short article: Goal relevance and artificial grammar learning. *Quarterly Journal of Experimental Psychology*, *62*(2), 228–238. https://doi.org/10.1080/17470210802479113

Ericsson, K. A., & Harwell, K. W. (2019a). Deliberate practice and proposed limits on the effects of practice on the acquisition of expert performance: Why the original definition matters and recommendations for future research. *Frontiers in Psychology*, *10*. https://www.frontiersin.org/article/10.3389/fpsyg.2019.02396

Ericsson, K. A., & Harwell, K. W. (2019b). Deliberate Practice and Proposed Limits on the Effects of Practice on the Acquisition of Expert Performance: Why the Original Definition Matters and Recommendations for Future Research. *Frontiers in Psychology*, *10*. https://www.frontiersin.org/article/10.3389/fpsyg.2019.02396

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). *The role of deliberate practice in the acquisition of expert performance*. *100*(3), 363–406.

Ettlinger, M., Margulis, E., & Wong, P. (2011). Implicit memory in music and language. *Frontiers in Psychology*, *2*. https://www.frontiersin.org/article/10.3389/fpsyg.2011.00211

Fitts, P., M., & Posner, M., I. (1967). *Human Performance* (Belmont:Brooks/Cole).

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

Forrin, N. D., & MacLeod, C. M. (2018). Contingency proportion systematically influences contingency learning. *Attention, Perception, & Psychophysics*, *80*(1), 155–165. https://doi.org/10.3758/s13414-017-1424-4

Geiser, E., Notter, M., & Gabrieli, J. D. E. (2012). A corticostriatal neural system enhances auditory perception through temporal context processing. *Journal of Neuroscience*, *32*(18), 6177–6182. https://doi.org/10.1523/JNEUROSCI.5153-11.2012

Geukes, S., Vorberg, D., & Zwitserlood, P. (2019). Disentangling semantic and response learning effects in color-word contingency learning. *PLOS ONE*, *14*(5), e0212714. https://doi.org/10.1371/journal.pone.0212714

Greenwald, A. G., Klinger, M. R., & Schuh, E. S. (1995). Activation by marginally perceptible ("subliminal") stimuli: Dissociation of unconscious from conscious cognition. *Journal of Experimental Psychology: General*, *124*(1), 22–42. https://doi.org/10.1037/0096-3445.124.1.22

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

Grégoire, L., Perruchet, P., & Poulin-Charronnat, B. (2014a). About the unidirectionality of interference: Insight from the musical Stroop effect. *The Quarterly Journal of Experimental Psychology*, *67*(11), 2071–2089. https://doi.org/10.1080/17470218.2014.896932

Grégoire, L., Perruchet, P., & Poulin-Charronnat, B. (2014b). Is the musical Stroop effect able to keep its promises? *Experimental Psychology*, *61*(1), 80–83. https://doi.org/10.1027/1618-3169/a000222

Grégoire, L., Perruchet, P., & Poulin-Charronnat, B. (2015). How does Stroop interference change with practice? A reappraisal from the music Stroop paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 417–425. https://doi.org/10.1037/xlm0000044

Grégoire, L., Poulin-Charronnat, P., & Perruchet, P. (2019). Stroop interference depends also on the level of automaticity of the to-be-interfered process. *Acta Psychologica*, *197*, 143–152. https://doi.org/10.1016/j.actpsy.2019.05.013

Greville, J., & Buehner, M. (2010). Temporal predictability facilitates causal learning. *Journal of Experimental Psychology. General*, *139*(4), 756–771.

https://doi.org/10.1037/a0020976

Greville, W. J., & Buehner, M. J. (2010). Temporal predictability facilitates causal learning. *Journal of Experimental Psychology: General*, *139*(4), 756–771. https://doi.org/10.1037/a0020976

Gromko, J. E. (2004). Predictors of music sight-reading ability in high school wind players. *Journal of Research in Music Education*, *52*(1), 6–15. https://doi.org/10.2307/3345521

Gudmundsdottir, H. R. (2010). Advances in music-reading research. *Music Education Research*, *12*(4), 331–338. https://doi.org/10.1080/14613808.2010.504809

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

Hardy, D. (1998). Teaching sight-reading at the piano: Methodology and significance. *Piano Pedagogy Forum*, *1*, 49–61.

Hargreaves, D. J. (1986). Developmental psychology and music education. *Psychology of Music*, *14*(2), 83–96. https://doi.org/10.1177/0305735686142001

Hayward, C. M., & Eastlund Gromko, J. (2009). Relationships among music sight-reading and technical proficiency, spatial visualization, and aural discrimination. *Journal of Research in Music Education*, *57*(1), 26–36. https://doi.org/10.1177/0022429409332677

Hébert, S., & Cuddy, L. L. (2006). Music-reading deficiencies and the brain. *Advances in Cognitive Psychology*, *2*(2–3), 199–206. https://doi.org/10.2478/v10053-008-0055-7

Heun, R., Burkart, M., Wolf, C., & Benkert, O. (1998). Effect of presentation rate on word list learning in patients with dementia of the alzheimer type. *Dementia and Geriatric Cognitive Disorders*, *9*(4), 214–218. https://doi.org/10.1159/000017049

Hirose, H., Kubota, M., Kimura, I., Ohsawa, M., Yumoto, M., & Sakakihara, Y. (2002). People with absolute pitch process tones with producing P300. *Neuroscience Letters*, *330*(3), 247–250. https://doi.org/10.1016/S0304-3940(02)00812-1

Honing, H., & Ploeger, A. (2012). Cognition and the evolution of music: Pitfalls and prospects. *Topics in Cognitive Science*, *4*(4), 513–524. https://doi.org/10.1111/j.1756-8765.2012.01210.x

How, E. R., Tan, L., & Miksza, P. (2021). A PRISMA review of research on music practice. *Musicae Scientiae*, 10298649211005532. https://doi.org/10.1177/10298649211005531

Iorio, C., Šaban, I., Poulin-Charronnat, B., & Schmidt, J. R. (2022a). EXPRESS: Incidental learning in music reading: The music contingency learning task. *Quarterly Journal of Experimental Psychology*, 17470218221092780. https://doi.org/10.1177/17470218221092779

Iorio, C., Šaban, I., Poulin-Charronnat, B., & Schmidt, J. R. (2022b). Incidental learning in music reading: The music contingency learning task. *Quarterly Journal of Experimental Psychology*. https://doi.org/10.1177/17470218221092779

JASP, T. (n.d.). *JASP (Version 0.11.1.0)*.

Jiang, Y., & Chun, M. M. (2001). Selective attention modulates implicit learning. *The Quarterly Journal of Experimental Psychology Section A*, *54*(4), 1105–1124. https://doi.org/10.1080/713756001

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*(3), 459–461.

Kamin, L., J. (1969). Predictability, surprise, attention and coding. In *Punishement and aversive behavior.* (Campbell, B. A., Church, R. M., pp. 279–296).

Kelly, M. H., & Martin, S. (1994). Domain-general abilities applied to domain-specific tasks: Sensitivity to probabilities in perception, cognition, and language. *Lingua*, *92*, 105–140. https://doi.org/10.1016/0024-3841(94)90339-5

Kerka, S. (2000). Incidental Learning. Trends and Issues Alert No. 18. *ERIC Publication*. https://eric.ed.gov/?id=ED446234

Kiss, M., Nemeth, D., & Janacsek, K. (2019). *Stimulus presentation rates affect performance but not the acquired knowledge – Evidence from procedural learning* [Preprint]. Neuroscience. https://doi.org/10.1101/650598

Kóbor, A., Kardos, Z., Horváth, K., Janacsek, K., Takács, Á., Csépe, V., & Nemeth, D. (2021). Implicit anticipation of probabilistic regularities: Larger CNV emerges for unpredictable events. *Neuropsychologia*, *156*, 107826. https://doi.org/10.1016/j.neuropsychologia.2021.107826

Kopiez, R., & In Lee, J. (2006). Towards a dynamic model of skills involved in sight reading music. *Music Education Research*, *8*(1), 97–120. https://doi.org/10.1080/14613800600570785

Kopiez, R., & In Lee, J. (2008). Towards a general model of skills involved in sight reading music. *Music Education Research*, *10*(1), 41–62. https://doi.org/10.1080/14613800701871363

Langan-Fox, J., Armstrong, K., Balvin, N., & Anglim, J. (2002). Process in skill scquisition: Motivation, interruptions, memory, affective states, and metacognition. *Australian Psychologist*, *37*(2), 104–117. https://doi.org/10.1080/00050060210001706746

Lange, K. (2009). Brain correlates of early auditory processing are attenuated by expectations for time and pitch. *Brain and Cognition*, *69*(1), 127–137. https://doi.org/10.1016/j.bandc.2008.06.004

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

Lehmann, A., C. (Ed.). (1997). The acquisition of expertise in music: Efficiency of deliberate practice as a moderating variable in accounting for sub-expert performance. In *Perception And Cognition Of Music* (0 ed., Vol. 1). Psychology Press. https://doi.org/10.4324/9780203344262

Lehmann, A., C., & Kopiez, R. (2009). Sight-reading. In *The Oxford handbook of music psychology* (pp. 344–351). Oxford: Oxford University Press.

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.

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., & 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

Lewis, P. A., & Miall, R. C. (2003). Distinct systems for automatic and cognitively controlled time measurement: Evidence from neuroimaging. *Current Opinion in Neurobiology*, *13*(2), 250–255. https://doi.org/10.1016/S0959-4388(03)00036-9

Lin, O. Y.-H., & MacLeod, C. M. (2018). The acquisition of simple associations as observed in color–word contingency learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(1), 99–106. https://doi.org/10.1037/xlm0000436

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*(4), 492. https://doi.org/10.1037/0033-295X.95.4.492

Logan, G. D., & Klapp, S. T. (1991). Automatizing alphabet arithmetic: Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(2), 179–195. https://doi.org/10.1037/0278-7393.17.2.179

Loui, P., Li, H. C., Hohmann, A., & Schlaug, G. (2011). Enhanced cortical connectivity in absolute pitch musicians: A model for local hyperconnectivity. *Journal of Cognitive Neuroscience*, *23*(4), 1015–1026. https://doi.org/10.1162/jocn.2010.21500

Loui, P., Wessel, D. L., & Kam, C. L. H. (2010). Humans rapidly learn grammatical structure in a new musical scale. *Music Perception*, *27*(5), 377–388. https://doi.org/10.1525/mp.2010.27.5.377

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

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

MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences*, *4*(10), 383–391. https://doi.org/10.1016/S1364-6613(00)01530-8

Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis.

Psychological Science, *25*(8), 1608–1618. https://doi.org/10.1177/0956797614535810

Matzel, L. D., Held, F. P., & Miller, R. R. (1988). Information and expression of simultaneous and backward associations: Implications for contiguity theory. *Learning and Motivation*, *19*(4), 317–344. https://doi.org/10.1016/0023-9690(88)90044-6

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

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

Mills, J., & McPherson, G. E. (2006). Musical literacy. In *The child as musician: A handbook of musical development* (Oxford University Press, pp. 155–172). Oxford: Oxford University Press.

Mishra, J. (2014a). Improving sightreading accuracy: A meta-analysis. *Psychology of Music*, *42*(2), 131–156. https://doi.org/10.1177/0305735612463770

Mishra, J. (2014b). 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

Montgomery, D. E., & Koeltzow, T. E. (2010). A review of the day–night task: The Stroop paradigm and interference control in young children. *Developmental Review*, *30*(3), 308–330. https://doi.org/10.1016/j.dr.2010.07.001

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*(2), 297–326. https://doi.org/10.1037/0033-2909.132.2.297

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

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). England: Oxford University Press.

Pearce, M., & Rohrmeier, M. (2012). Music cognition and the cognitive sciences. *Topics in*

Cognitive Science, *4*(4), 468–484. https://doi.org/10.1111/j.1756-8765.2012.01226.x

Perea, M., García-Chamorro, C., Centelles, A., & Jiménez, M. (2013). Position coding effects in a 2D scenario: The case of musical notation. *Acta Psychologica*, *143*(3), 292–297. https://doi.org/10.1016/j.actpsy.2013.04.014

Perruchet, P. (2019). What mechanisms underlie implicit statistical learning? Transitional probabilities versus chunks in language learning. *Topics in Cognitive Science*, *11*(3), 520–535. https://doi.org/10.1111/tops.12403

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches—ScienceDirect. *Trends in Cognitive Sciences*, *10*(5), 233–238.

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.

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

Rohrmeier, M., & Cross, I. (2009). *Tacit tonality: Implicit learning of context-free harmonic structure*. 10.

Rohrmeier, M., & Rebuschat, P. (2012). Implicit learning and acquisation of music. *Topics in Cognitive Science*, *4*(4), 525–553. https://doi.org/10.1111/j.1756-8765.2012.01223.x

Rohwer, D., & Polk, J. (2006). Practice behaviors of eighth-grade instrumental musicians. *Journal of Research in Music Education*, *54*, 350–362. https://doi.org/10.1177/002242940605400407

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

Saariluoma, P. (1994). Location coding in chess. *The Quarterly Journal of Experimental Psychology Section A*, *47*(3), 607–630. https://doi.org/10.1080/14640749408401130

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

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

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*(4), 606–621. https://doi.org/10.1006/jmla.1996.0032

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

Schmidt, J. R. (2012). Human Contingency Learning. *Encyclopedia of the Sciences of Learning, Spinger, Boston,MA.*, 1455–1456. https://doi.org/10.1007/978-1-4419-1428-6_646

Schmidt, J. R. (2018). Best not to bet on the horserace: A comment on Forrin and MacLeod (2017) and a relevant stimulus-response compatibility view of colour-word contingency learning asymmetries. *Memory & Cognition*, *46*(2), 326–335. https://doi.org/10.3758/s13421-017-0755-7

Schmidt, J. R. (2021a). Apprentissage incident des associations simples de stimulus-réponse: Revue de la recherche avec la tâche d'apprentissage de contingences couleur-mot: *L'Année psychologique*, *Vol. 121*(2), 77–127. https://doi.org/10.3917/anpsy1.212.0077

Schmidt, J. R. (2021b). Incidental learning of simple stimulus-response associations: A review of color-word contingency learning research. *LAnnee psychologique*, *121*(2), 77–127.

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). 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. (2012b). Does temporal contiguity moderate contingency learning in a speeded performance task? *Quarterly Journal of Experimental Psychology*, *65*(3), 408–425. https://doi.org/10.1080/17470218.2011.632486

Schmidt, J. R., & De Houwer, J. (2012c). 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. (2012d). 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. (2016). Time course of colour-word contingency learning: Practice curves, pre-exposure benefits, unlearning, and relearning. *Learning and Motivation*, *56*, 15–30. https://doi.org/10.1016/j.lmot.2016.09.002

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., & De Houwer, J. (03/2016a). Contingency learning tracks with stimulus-response proportion: No evidence of misprediction costs. *Experimental Psychology*, *63*(2), 79–88. https://doi.org/10.1027/1618-3169/a000313

Schmidt, J. R., De Houwer, J., & Besner, D. (2010). Contingency learning and unlearning in the blink of an eye: A resource dependent process. *Consciousness and Cognition*, *19*(1), 235–250. https://doi.org/10.1016/j.concog.2009.12.016

Schmidt-Kassow, M., Schubotz, R. I., & Kotz, S. A. (2009). Attention and entrainment: P3b varies as a function of temporal predictability. *NeuroReport*, *20*(1), 31–36. https://doi.org/10.1097/WNR.0b013e32831b4287

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic Human information processing: I. detection, search, and attention. *Psychological Review*, *84*, 127–190.

Schön, D., Anton, J. L., Roth, M., & Besson, M. (2002). An fMRI study of music sight-reading. *NeuroReport*, *13*(17), 2285–2289. https://doi.org/10.1097/01.wnr.0000044224.79663.f5

Schön, D., Semenza, C., & Denes, G. (2001). Naming of musical notes: A selective deficit in one musical clef. *Cortex*, *37*(3), 407–421. https://doi.org/10.1016/S0010-9452(08)70581-1

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

Schwartze, M., Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2011). Temporal regularity effects on pre-attentive and attentive processing of deviance. *Biological Psychology*, *87*(1), 146–151. https://doi.org/10.1016/j.biopsycho.2011.02.021

Scripp, L. R. (1995). *The development of skill in reading music* [Thesis dissertation]. Harvard University,MA.

Selchenkova, T., Jones, M. R., & Tillmann, B. (2014). The influence of temporal regularities on the implicit learning of pitch structures. *Quarterly Journal of Experimental Psychology*, *67*(12), 2360–2380. https://doi.org/10.1080/17470218.2014.929155

Sergent, J., Zuck, E., Terriah, S., & MacDonald, B. (1992). Distributed neural network underlying musical sight-reading and keyboard performance. *Science*, *257*(5066), 106–109. https://doi.org/10.1126/science.1621084

Shahin, A., Bosnyak, D. J., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *Journal of Neuroscience*, *23*(13), 5545–5552. https://doi.org/10.1523/JNEUROSCI.23-13-05545.2003

Shanks, D. R. (2007). Associationism and cognition: Human contingency learning at 25. *Quarterly Journal of Experimental Psychology*, *60*(3), 291–309. https://doi.org/10.1080/17470210601000581

Sloboda, J. A. (1976). Visual perception of musical notation: Registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology*, *28*(1), 1–16. https://doi.org/10.1080/14640747608400532

Sloboda, J. A. (1982). Music performance. In *The psychology of music* (pp. 479–496). Academic Press.

Sloboda, J. A. (1985). *The musical mind: The cognitive psychology of music*.

Sloboda, J. A. (2000). Individual differences in music performance. *Trends in Cognitive Sciences*, *4*(10), 397–403. https://doi.org/10.1016/S1364-6613(00)01531-X

Stanzione, M., Grossi, D., & Roberto, L. (1990). Note-by-note music reading: A musician with letter-by-letter reading. *Music Perception*, *7*(3), 273–284. https://doi.org/10.2307/40285465

Staresina, B. P., & Davachi, L. (2009). Mind the gap: Binding experiences across space and time in the human hippocampus. *Neuron*, *63*(2), 267–276. https://doi.org/10.1016/j.neuron.2009.06.024

Stewart, L. (2005). A neurocognitive approach to music reading. *Annals of the New York Academy of Sciences*, *1060*(1), 377–386. https://doi.org/10.1196/annals.1360.032

Stewart, L., Henson, R., Kampe, K., Walsh, V., Turner, R., & Frith, U. (2003). Brain changes after learning to read and play music. *NeuroImage*, *20*(1), 71–83. https://doi.org/10.1016/S1053-8119(03)00248-9

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

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643.

Takeuchi, A., H., & Hulse, S., H. (1993). Absolute pitch. *Psychological Bulletin*, *113*(2), 345.

Theusch, E., & Gitschier, J. (2011). AbsolutepPitch twin study and segregation analysis. *Twin Research and Human Genetics*, *14*(2), 173–178. https://doi.org/10.1375/twin.14.2.173

Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological Bulletin*, *139*(4), 792–814. https://doi.org/10.1037/a0030801

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 (Second Edition)* (pp. 265–298). Academic Press. https://doi.org/10.1016/B978-012213564-4/50009-3

Waters, A. J., Underwood, G., & Findlay, J. M. (1997). Studying expertise in music reading: Use of a pattern-matching paradigm. *Perception & Psychophysics*, *59*(4), 477–488. https://doi.org/10.3758/BF03211857

Wlotko, E. W., & Federmeier, K. D. (2015). Time for prediction? The effect of presentation rate on predictive sentence comprehension during word-by-word reading. *Cortex*, *68*, 20–32. https://doi.org/10.1016/j.cortex.2015.03.014

Wolf, T. (1976). A cognitive model of musical sight-reading. *Journal of Psycholinguistic*

Research, *5*(2), 143–171. https://doi.org/10.1007/BF01067255

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

Zacks, R. T., & Hasher, L. (2002). Frequency processing: A twenty-five year perspective. In P. Sedlmeier & T. Betsch (Eds.), *Etc. Frequency Processing and Cognition* (pp. 21–36). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198508632.003.0002

Zakay, D., & Glicksohn, J. (1985). Stimulus congruity and S-R compatibility as determinants of interference in a Stroop-like task. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *39*(3), 414–423. https://doi.org/10.1037/h0080069

Zatorre, R. J. (2003). Absolute pitch: A model for understanding the influence of genes and development on neural and cognitive function. *Nature Neuroscience*, *6*(7), 692–695. https://doi.org/10.1038/nn1085

Zatorre, R. J., Perry, D. W., Beckett, C. A., Westbury, C. F., & Evans, A. C. (1998). Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. *Proceedings of the National Academy of Sciences*, *95*(6), 3172–3177. https://doi.org/10.1073/pnas.95.6.3172