Cross-Modal Schema Effect of Music Pairing on Shape Sequence Acquisition

A Dissertation
Presented to
The Academic Faculty

by

Yiren Ren

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Psychology

Georgia Institute of Technology
August 2021

Cross-Modal Schema Effect of Music Pairing on Shape Sequence Acquisition

Approved by:

Dr. Thackery Brown, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Audrey Duarte
School of Psychology
Georgia Institute of Technology

Dr. Grace Leslie
School of Music
Georgia Institute of Technology

Date Approved: [July 5, 2021]

ACKNOWLEDGEMENTS

I appreciated all undergraduate students across Georgia Institute of Technology community who participated in this cognitive experiment. The study would not complete without these volunteers' efforts and helps.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Music is a multidimensional sequence of pitches and temporal intervals that has a predictable structure over time. Prior literature has revealed that humans are innately equipped to learn and anticipate these pitches and intervals. Because of the important role music plays in humans' daily lives, learning how music interacts with other cognitive processes would help future utilization of music in clinical or applied ways. Though many studies have tested how to use music as a tool to improve other cognitive functions, fewer studies have investigated how music potentially affects memory encoding for information other than the music itself. This question is worth investigating, given music’s regular structure and frequent presence in the background while we study or simply experience our daily lives. Schema theory has shown that new information that is related to a learned memory structure can be encoded and learned faster, although this has never been directly tested, to the best of our knowledge, in the context of the learned structure being music. Thus, this study aimed to apply schema theory using an association between musical sequence properties and the workload required for parallel visual item sequence encoding – in doing so, I tested whether listening to familiar and regular music provided a "temporal schema" through its organized and hierarchical structure that has a cross-modal influence on the acquisition of other (here: visual) mnemonic information. Consistent with my hypotheses, the results revealed an interactive effect of music familiarity and music regularity on parallel visual sequential learning. While listening to music may improve or distract from parallel memory encoding in various circumstances, this study not only provided novel evidence that music regularity and familiarity were both factors

determining the music's influence, but also implied that music's effect on memory might depend on other individual differences factors.

CHAPTER 1. INTRODUCTION

Can you recall the very last time when you heard a piece of music? It was usually not a remote event for most of us. According to a recent survey (2017), Americans on average listened to music for more than 32 hours within a week. Music has been deeply ingrained in human daily life from the early stages of human history (Cross, 2001; Huron, 2001). Most people find music enjoyable, and humans have remarkable, innate music acquisition and perception abilities (Menon & Levitin, 2005; Norton et al., 2005; Patel, 2010). Because of this, nowadays with easier access to music such as via Bluetooth, it is frequent to see people listening to music while doing other things, including when they perform important tasks like studying or working. This phenomenon has led researchers to question what effects music may have on other cognitive functions. For example, studies have provided evidence for facilitative effects of listening to music during motor tasks in patients afflicted with Morbus Parkinson (Bernatzky et al., 2004). Other studies have focused on how music modulated emotion and have demonstrated the possibility of using music as an intervention in mood disorder treatments (Moore, 2013).

However, regarding its effect on memory, there has been debate over whether background music helps or distracts from parallel memory-related tasks (Furnham & Bradley, 1997; Jäncke & Sandmann, 2010). Some evidence suggested that listening to music could disrupt attention to working memory tasks such as reading comprehension, stealing cognitive resources and resulting in worse memory results (Furnham & Strbac, 2002; Thompson et al., 2012). Others have found that using music paired with information could help memory performance such as improving autobiographical memory recall in

Alzheimer's disease patients by providing contextual cues (Cuddy et al., 2017; Lord & Garner, 1993) as well as verbal memory and language learning (Kang & Williamson, 2014; Smith, 1985).

There is therefore a long history of researchers testing music as a memory aid but the results have been different. Moreover, many of these studies have focused on verbal/linguistic related memory as well, but from an applied research angle this is not the only type of memory to suffer in neuropsychological conditions. The answer to whether music could help people memorize information, thus remains unclear and controversial, especially for non-auditory and non-linguistic information. There are a number of possible explanations for this continued uncertainty. Firstly, not all music is created equal (meaning different genres, and individual compositions within a genre have distinctive properties that could influence their impact on memory) and not all people are equally good at processing music (Jonides, 2008; Negus, 2013; Seashore, 1937). <u>This led me to consider a fundamental question in this thesis</u>: does music of different properties (e.g. more- versus less-regular, familiar versus unfamiliar, etc.) have different effects on non-music information memory encoding? Moreover, music's effect on memory and learning might be inconsistent due to people's different music skills.

Another possible explanation for the undetermined effect of music on memory is that music might show unequal effects on different types of memory (e.g., active working memory vs. long-term episodic memory vs. semantic memory) or even for different types of target information being memorized. For example, many studies suggesting benefits of background music for memory encoding used text information as the memory target – this may be a critical detail because the overlap in involvement between our language system and music system has been implied as one of the explanations of such facilitative music pairing effects (Ferreri et al., 2013; Moussard et al., 2012; Patel, 2010). Meanwhile, less evidence has been found to clarify how background music influences learning non-text/non-verbal/non-linguistic information, and it is further the case that there is a lack of research focused on music pairing just during the memory encoding stage. From an applied angle that is an important limitation of the current literature, because it may not always be feasible to plan on replaying music from encoding to elicit retrieval in an individual (e.g., Cuddy et al., 2017; Lord & Garner, 1993). Considering the prevalence of music listening in daily life during information processing and encoding, in this study I sought to fill the gap on whether listening to background music influences non-text and non-auditory information processing during memory encoding. Due to the variability of memory types, information categories, and music features in the existing literature, being able to answer the question of for whom (individual differences), when (e.g., its effect during memory recall vs. memory encoding), and what type of music could improve memory encoding will ultimately require more than one research experiment. But any more precise insights or discoveries about these questions would be important from both theoretical and applied interests. The answers could help us understand more about the basic science of how cross-modal memory works, and could further the applied science of helping humans modulate their learning and working environment and addressing the idea that we could perhaps use music to optimize the efficiency of cognitive functions. For instance, as a student, I often listen to music while preparing for an exam – do any properties of the music matter for whether it is beneficial? In this study I aimed to narrow down the broad question of how music (broadly defined) affects memory (broadly addressed in the prior literature) to focus on how background music affects visual sequential memory, which on one hand is a type of memory that does not share as much cognitive and neural substrate as linguistics and music do, but on the other hand is a major component of daily episodic memory and associations, as well as other long-term memory important for daily function like being able to retrieve a sequence of visual landmarks to navigate our environment. I further aimed to make the broad research topic manageable taking the concern raised above about potential variability in the prior literature in music genres and specific compositions used as stimuli, and zooming in on manipulating properties of music that can be more directly related to the memory sequencing psychology literature. By asking such questions about a specific memory type and manipulating specific properties of music, one of my goals was to better identify whether a specific constructions of music could improve long-term sequence memory encoding.

With this goal of studying specific mechanisms that modulate how music and other memories interact, in this thesis, I raised a new theoretical perspective for the music memory field on how music could be used to help sequential associative memory formation in other modalities. I proposed that research area can be moved forward by building a bridge between the knowledge that our minds have a schematic music perception system and psychological theory surrounding "schemas" in human associative memory. The psychology of memory suggests that schemas (learned and stable associative memory structures/networks) enable new associated information to be encoded easier (Van Kesteren et al., 2012). Music cognitive science has indicated that that human music processing system could be connected to psychological schema because music perception was highly depended on prior knowledges in a way that music can become quickly schematized and that learned music, which follows music syntax, could act as schemas to help new music learning (Leman, 2012). Based on these two theoretical bases (which were elaborated in further detail in the following text), I hypothesized that schematic music sequences could help new sequence encoding for associated information . By building a design in which I compose and manipulate the schematic nature of music stimuli, and pair it with visual sequences of information, the following thesis project aimed to reveal information that could inform strategies for pairing music that could potentially aid encoding of other memory sequences in our daily life (e.g., the sequence of sights contributing to an episodic memory or a navigational route, which can later – with multiple exposures - contribute forming or updating semantic knowledge).

One important clue for my idea was the evidence *at retrieval* from both healthy individuals and those with memory-deficits that music provided sequential cues for memory and helped memory for associated information, putatively by providing reinstatement of associated temporal contexts (Kasdan & Kiran, 2018; Samson & Zatorre, 1992). However, as noted above, there has been debate on the *encoding* side of memory over whether music played in the background of parallel information could positively affect learning results – there has been some evidence that listening to music during encoding could improve episodic memory performance (Ferreri et al., 2013; Proverbio et al., 2015), but music has also shown to impair visual associative memory performance in older adults (Reaves et al., 2016). In either case, the mechanisms for why music exerts such influences have also been unclear. Thus, one open question to be addressed was why music does not persistently benefit encoding across people and what properties of music matter for its influence.

Music is a highly complex stimulus. To constrain the research problem from the natural variability in music stimuli one might take from the radio to something that can be carefully manipulated in the laboratory, in this thesis I focused on two properties of music – familiarity and regularity. As noted above, music science has indicated music could form "temporal schemas" (Leman, 2012) – encoded sequential frameworks of pitches and timing that enabled people to predict what comes next in the composition with relatively limited effort and attention. As such, I hypothesized that one implication from theories of *non-temporal* schemas (e.g., Van Kesteren et al., 2010) could be applied to music: Namely, in non-musical associative memory, new learning that is congruent with the structure of existing schemas is encoded and consolidated easier and faster, putatively because existing associations help bridge the new associations to be formed (Van Kesteren et al., 2010). For example, it may be easier to learn new locations on campus when other information about the structure of the environment is already known (Van Kesteren et al., 2018). If this theory translates to temporal schemas, then my hypothesis was that the regular and temporal structure of a piece of familiar music could provide a sequential "scaffold" through which new information – encoded in parallel - could become associated in sequence. Because I theorized the benefit came from learned (familiar) and regular music providing an existing predictive sequential structure, I predicted further that the benefit should not be limited to sequencing new auditory stimuli – rather, schema theory could provide a way through which background music might provide a rhythm and order to sequencing new information more generally, and thus help people encode and sequence arbitrary stimuli from various modalities in our lives.

To test my theory, I therefore paired novel, arbitrary visual shape sequences with music during a visual sequential learning task. Here in this study, I defined music as a sequence of tones ordered in a specific temporal relationship. Based on the schema theory, new memory encoding can benefit from being able to relate it to a familiar memory structure (Van Kesteren et al., 2012). Thus, <u>**my first primary hypothesis** for my thesis was that visual memory sequence encoding could benefit more from pairing with known music sequences more than novel ones where the structure is unknown</u>. In the experiment setting, I set up two conditions of **music familiarity (first independent variable) –** learned music and unlearned music (more details in Method and Results).

Besides familiarity of music, I also hypothesized that the **predictivity or regularity (second independent variable)** of music would also have different effects on paired visual sequence learning, motivated by evidence in psychological schema theory that a stable and expected old memory structure (here a music sequence) could provide more information for structuring new long-term memories (here sequence associations for visual stimuli). Music science has found that music, similar to language, followed rules and syntax such as chords and octaves (Patel, 2003; Tan et al., 2017). In this study design, I aimed to manipulate music in the level of its syntactical correctness and referred it as music regularity. Admittedly, "Regularity" was not a technically-specific term in music science –I defined it by how I operationalized this term below as well as in the Method. In brief, I composed novel music stimuli following a specific predictive syntax and then created their comparable condition by breaking the syntactic rules while the musical components (notes) remained the same in order to ensure the two conditions only differed in their predictivity (see following and *Method*. for details). At a high level, the "regularity" manipulation I used was based on evidence that "syntactically correct" music could form schemas for new music perception and learning (Leman, 2012). Something not tested before: here I aimed to leverage this evidence to investigate if such schematic music (music that following musical rules) could facilitate non-music sequence encoding. Similar to language, not all music genres are structured in the same manner, and not all individual music pieces are composed equally. What I targeted in this thesis was the idea that music has different levels of "regularity" – I used the terminology that music was "regular" when the arrangement of tones in temporal order was organized and statistically associated with each other (in music terms, the notes follow musical syntax and are linearly arranged). Studies showed that such regular music is predictable even for the first time listening, and such music evokes brain activity differently from syntactically irregular music with expectation violations (Hagoort, 2003; Koelsch et al., 2007, 2019; Patel et al., 1998). Such data implied that humans are equipped with predictive coding for syntax-correct music. Thus, by manipulating regularity in experimentally-generated music, rather than allowing to be a freely varying parameter that might have contributed to variability in the prior literature on how music affects different types of memory, I could test whether the predictive/regular music affected parallel sequence learning differently than irregular music, as predicted by psychological schema theory.

A highly relevant concept to my operationalization of "Regularity" is prior studies on human recognition of music with more or less regularity (represented by comparing tonal vs. atonal music) also have shown that humans could fail to perceive syntactically incorrect music as an event with hierarchical structure (Dibben, 1994). This hierarchy in music is important, because it means that notes are perceived as grouped into motifs, which are then grouped into lines or sections, which are grouped into movements and stanzas (Lerdahl & Jackendoff, 1983). A non-music analog is that in language sentences are made up by phonemes grouped into words grouped into phrases - and it turns out that regularly-structured music's hierarchy enables human to better encode and understand the music (Levitin, 2006), and therefore this hierarchical structure and relationships of pitch and temporal intervals are important for music to be a meaningful event (and something that is degraded in irregular music) (Bigand et al., 2014; Dibben, 1994). To relate this back to my thesis idea, based on schema theory, if schemas could facilitate information encoding by providing a persistent template for how associations are organized (van Kesteren et al., 2014), and given evidence from music science that that applies to temporal dimension as well (i.e., for music's sequential associations), then irregularly-structured music – with low level of hierarchical structure and looser relationship between pitches and temporal intervals - should provide a less-effective sequential schema for new learning. <u>Thus **my second primary hypothesis for my thesis** was that irregular music would thus result in less benefit or even impaired paired long-term sequence information encoding, compared to regular music</u>.

To test this second hypothesis, I created purely-instrumental music stimuli, following western classical music composition, which represented the type of music people might listen to during their daily life. I manipulated this classically-composed music into two conditions of music based on the regularity. In the study, my "regular" music composition stimuli, or syntactically schematic music, followed the predictive rules of western music. For example, it was regular music when keys were arranged by a flow of perfect fifth. In contrast, the irregular music condition lost/violated the regular predictable pattern of tones and temporal interval, which as I described in detail above was predicted as a necessary feature for the music stimuli to provide sequential schemas for new visual sequences learning (i.e., for how classical music on the radio might help us encode the series of landmarks we pass during navigation). Importantly, however, in my design irregular music had the same set of tones across the composition as a whole as regular music so that varying tones or keys were not a possible confound for comparisons between sequence learning paired with my regular vs irregular stimuli, and so that irregular music was still memorable such that people were able to get familiar with it (more details in Method), ensuring a clean contrast based on regularity. To avoid possible confusion from the music-related terminology, in the following text when I referred to either of my musical conditions generally (since both conditions are sequences made up of tones) I would call them simply “musical sequences” or “music”.

Although cross-modal schema effects have not been directly tested in the manner proposed here, there have been hints from neuroscience studies that further support my predictions for their possible influences in my study: First of all, at a high level my study aimed to study cross-modal schema benefits on learning. We did know at least that the associations within schemas themselves could be multi-modal, involving abstract or higher-level knowledge and concepts in both humans and rodents. For example, schemas in school learning material could help encode new knowledge, which could be highly abstract and involves representations acquired through both text and oral/visual presentation (van Kesteren et al., 2014). Animals studies have shown that in the hippocampus, an area strongly associated with memory encoding and consolidation, schemas and associative memory were represented by hierarchically-organized neural representations, suggesting that relevant memories shared overlapping neural states even when the component pieces of associative information were of different types (e.g. item and spatial location) (McKenzie et al., 2014). Second of all, in this study I proposed that music would help encoding of sequences of new information by providing a stable and regular temporal order. Neuroscience has identified regions such as the hippocampus at not only being good at associating new cross-modal information (Gilbert & Kesner, 2002; Henke et al., 1999; Holdstock et al., 2010) but being a brain region with a remarkable ability to process temporal and sequential structure (Barnett et al., 2014; Eichenbaum, 2014; Kesner et al., 2002). Because such literature indicated shared neural substrate for associative memory development and temporal sequence encoding, for the purposes of my study this suggested the brain had mechanisms in place that could enable us to take advantage of temporal sequence representations that were provided by familiar music in order to reinforce the encoding of new information via such temporal structure.

Some other hints that my hypothesized effects of music on learning should occur came from studies of parallel-sequence learning: one study showed that learning could occur in parallel for two independent sequences when one of them was non-attentionally-demanding (as familiar music can be)(Curran & Keele, 1993). Second, when two streams of information were present (e.g. color and text), implicit sequence learning of either stream could happen only when the two streams of information were stably correlated, meaning that same pairs of individual stimuli of the sequences co-occur during learning (Weiermann et al., 2010). Such data further motivated my hypothesis that by building a pairing between music and visual sequences, people might benefit from implicit temporal sequence processing of the music in relation to the visual stimuli and improve their explicit visual sequence memory. Critically, however, they also suggested that if music was disorganized and unpredictable in its structure (as I tested through the irregular condition in the following design), this could produce 1) distracting effects of music (consistent with the interpretations from Furnham & Bradley, 1997; Jäncke & Sandmann, 2010) and 2) further impair encoding of new, parallel sequences of information because the sequential relationships in the music were themselves irregular.

For this thesis, I was particularly interested in the applied potential of this work, and the fact that it could help us understand alternative learning and memory strategies that may be applied to helping patients with memory deficits such as dementia, as exhibited in Alzheimer's disease, or even attention deficits. One exciting possibility was that retrieving sequential knowledge in other attention-free settings may also influence incidental learning of new information. For example, sequential and rhythmic body movements like dance might also be able to benefit new learning when paired. Critically, however, any applied angle for such work must consider individual differences. People could differ dramatically in their memory and their music abilities, and because of this, music interventions could differentially impact musicians vs non-musicians or simply any individual as a function of their innate music abilities. Because of this, this thesis also examined at the role of individual differences on my hypothesized cross-modal learning effects. The results might help our understanding of whether music modulates memory in some more than others. Given that most of our lives involve listening to music, at both the group and individual differences level, my study could provide evidence for *how* this practice influences our daily tasks – perhaps without us knowing it.

CHAPTER 2. BACKGROUND

2.1 Humans Use Music Schemas to Process and Anticipate Music.

Music has played an essential role in entertainment and communication across human history, and humans have evolved in a way that makes them highly adept at music recognition and production. Although a piece of music can contain complex information from multiple domains, humans can memorize it rapidly with small cognitive workload (Tan et al., 2017).

One of the skills that supports humans' music processing efficiency is the anticipation for both pitch sequences and temporal intervals making up the music "flow". The ability to predict future events is an essential function for animals to have faster responses to environmental changes, and associate negative (as well as positive) outcomes with sequences of events, and we see an interesting reflection of this in our processing of music. Such sequential prediction and learning processes have been tied in human neuroscience work to evidence for universal neural mechanisms that process error detection such as ERN (error-related negativity), a brain signal that occurs when the brain detects a behavioral error (M Falkenstein, 1990). Brain regions such as basal ganglia, medial prefrontal cortex and medial temporal lobe also support error detection and event prediction (Alexander & Brown, 2014; Michael Falkenstein et al., 2001; Knutson et al., 2003). Although such mechanism-level insight has been limited in studies of music, these neural circuits have been broadly implicated in studies of music processing and prediction. Many EEG studies have shown that people predict musical beats to follow a preferred rhythm, and the frequency of timing intervals in the music synchronize with oscillations in various brain areas (Large et al., 2015; Nakamura et al., 1999). One of the most frequently highlighted findings is that the human brain reacts to a violations of these predictions in what will come next in the music with event-related potentials (ERPs) such as the early right anterior negativity (ERAN) and frequency mismatch negativity (MMN) (Herholz et al., 2009). For example, ERAN usually happens 200ms after a chord that people "feel uncomfortable with" and MMN (fronto-central negativity) reacts most strongly for out-of-key tones (M. A. Rohrmeier & Koelsch, 2012) – both responses are ultimately thought to depend on having had prediction, from a schema of musical structure, for what chords and tones "should have come next". Other areas such as the auditory cortex and inferior frontal cortex have also been shown to support music prediction and music memory (Salimpoor et al., 2013; Watanabe et al., 2008).

Interestingly, the brain produces error-responsive potentials during both music and language processing that are similar, arguably because they both follow have basic syntax and structural (schematic) rules (Bigand et al., 2014). Interestingly, music syntax is not just found in humans. For example, primates can recognize octaves, which indicates that specific music rules exist across species and can be innate.

The syntax and structure are important, because studies have also shown that humans are better at processing musical sequences where durations of the temporal intervals between notes are in a ratio of specific integers (Janata & Grafton, 2003). Other studies have shown that people prefer to group tones with short intervals into one section or an musical episode (Tan et al., 2017). Although musical rhythm structure and styles change across cultures, humans generally expect regular rhythmic sequence patterns (Tan et al., 2017). These syntax or basic rules in music enable humans to predict ongoing music with good accuracy. One implication for the present proposal is that since people powerfully process the regularity of music and the predictability of music, and likewise music that is irregularly structured and scrambled is difficult to predict and process and this may influence parallel cognitive functions.

The syntax of music also helps humans to process even novel pieces of music in clusters, within which tones are harmonic to each other. The existence of chords is one example of defining clusters. Similar to visual perception, humans group auditory stimuli based on Gestalt or related rules (Tan et al., 2017). The human auditory system tends to detect well-regulated or prior learned patterns of input from both tonal and temporal dimensions, and is less efficient in learning irregular patterns of sound like noise (Tan et al., 2017). Integrating ongoing dynamics of sounds into clusters enables people to make chunks in working memory (called "phrases" in music), during listening that enable easier encoding. Collectively, these features are highly reminiscent of characteristics of other forms of "schemas" (discussed below) – with familiarity and statistical regularities in the structure of music enabling us to process and predict tones and timing efficiently, even the first time we hear a composition.

Behavioral studies have confirmed that music with tonal sequence patterns can be rapidly learned by mere exposure (M. Rohrmeier & Rebuschat, 2012). Besides the auditory system's rapid detection of the agreement of sounds with basic music rules and syntax, the efficient music learning system also heavily relies on encultured schemas, long-term memory for prior music experiences, and short-term repetition. Leman (2012) described in detail in his book how and why memory networks involved in schemas are also involved in processing music. In brief, he argued that when listening to one piece of novel music, listeners will form expectations based on musical patterns stored in memory and rapidly scan the architecture of the music by integration and segmentation. The outcomes of this procedure may be buffered and processed within the working memory network (often in chunks, as I noted above). Meanwhile, prior knowledge of music will help listeners to recognize familiar or preferred patterns and these patterns could be culture-specific. This detection, if any, can evoke and stimulate encoding processes in long-term memory networks. By automatically comparing the new auditory sequence with prior music memory, the new information can be more quickly stored due to an existed exemplar. The whole process, after being repeating with just a few exposures, can result in forming accurate cognitive representation for the newly learned music and its overlap with the old music memory, together clustering the features of new music with prior music knowledge into an updated version of the schema (Leman, 2012). In this way, our semantic knowledge of music can grow and evolve through the same principles attributed to other forms of knowledge in the memory literature.

In conclusion, the extant studies suggested music acquisition and recognition abilities were highly developed in humans. The mechanisms behind it involved several neural systems and a strong musical schematic processing system, putatively through the same schema processing circuits attributed to other forms of memory, enabled people to learn music unintendedly. People can memorize or even reproduce a piece of music without declarative understanding of the meaning or the grammar behind the music (Ettlinger et al., 2011). Because of the extraordinary music perception and the fact that many people love to listen to music while focusing on other executive function-demanding tasks, the important question this study asked was how background music affects memory for new knowledge learning and episodic encoding.

2.2 Can music enhance memory of other types?

Music is often used as mnemonic device for declarative memory. For example, many people learned the alphabet using a song when they were kids. Similarly, music can be a powerful memory cue, such that people might recall details and emotions from a remote event when hearing a piece of music again after a long time. Music can serve as a sequentially-structured delivery of temporal context for episodic memory and can be used to cue recall of an event (Belfi et al., 2016; Janata, 2009; Wallace, 1994). It is thus natural to consider the possibility and the mechanisms through which sequentially-structured music can aid human memory.

Some studies have shown that pairing music with text-type information could improve retrieval success when music is later played (Moussard et al., 2012; Rainey & Larsen, 2002). The ability of music to serve as a retrieval cue may be unsurprising – if each note, or perhaps a collection of notes, is associated with a stimulus and treated as an associative retrieval cue, this could support pattern completion of a memory. But does it aid encoding, per se? Several studies have shown that word spelling accuracy significantly increased when, during learning, the words were presented through music instead of by speech (Schlaug et al., 2010). Many commercial advertisements have taken advantage of this phenomenon and paired brief and easily memorized music to their slogans (Yalch, 1991). The pairing effect has also led to new interventions in some memory disorders. For example, one study showed that aphasia patients benefited from using melodic intonation therapy, which asked patients to sing instead of speaking out words (Albert et al., 1973). Similar therapy was used in Alzheimer's patients – their autobiographical memory retrieval has been shown to be better during listening to familiar music, which putatively provides contextual cues (Irish et al., 2006). Pairing music with visual cues such as film clips can help to recall of the visual stimuli (Boltz et al., 1991).

Importantly, most of the above current evidence for music's enhancement on non-music memory focuses on pairing music with text or only on how music, especially old-memory related and previously-experienced music - as an event context - can help evoke recall of the memory. There is a need for work that explicitly tests how music, both familiar and novel, might affect memory during encoding and learning. In the study, I tested whether familiar and regular music conveyed a benefit to learning associations between other stimuli (in my design, simple abstract shapes) during encoding, and whether it did so through the lens of the schema memory theory.

2.3 Information encoding and consolidation benefits from associating to the schema.

The above literature review focused on music properties and music schemas. What do we know about schemas from the broader memory literature? Standard consolidation theory holds that memory consolidation processes involve transferring instances of declarative memories formed in the hippocampus gradually, over time, out to the cortex – most information needs numerous revisits or internal replay events to become stable, lasting, hippocampus-independent memory (Gais & Born, 2004; Hasselmo, 2006; Nadel et al., 2000). However, striking data have shown that when new associative information is congruent in some way with the structure of previously encoded knowledge or old memories, referred to as a “schema”, the memory encoding process is more efficient and the consolidation process can be much faster (Tse et al., 2007; Van Kesteren et al., 2010). For example, the first time a child sees an eagle, he/she might have no idea what that animal is. But with prior knowledge that birds have wings and beaks, he/she could easily encode the species of eagles, categorize this type of animal with birds rather than - e.g. - with dogs, and make predictions about other characteristics of eagles that have not yet been observed.

The schema can be highly abstract and consist of hierarchical information – it encompasses different dimensions of the memories and thus may be multi-modal (McKenzie et al., 2014). Typically, schemas have been studied in non-temporal forms (e.g., knowledge of spatial environment structure or the relationships between concepts), but the critical finding is that novel associations (e.g., that eagles often fish) are better learned when present in a familiar context or semantic framework than an unexpected one (e.g., when other behaviors and features of birds are known than when eagles are the first bird ever encountered or learned about (Van Kesteren et al., 2010). Evidence from cognitive neuroscience has linked several neural mechanisms to the learning benefit of the schema (McKenzie et al., 2013; Preston & Eichenbaum, 2013; Van Kesteren et al., 2010, 2018). It has been argued that an interaction between the medial prefrontal cortex, which helps survey old memory traces and detect matches/mismatches with the new memories, and the medial temporal lobe - a core region for episodic memory encoding – enables congruent information with prior knowledge to fast-track strong encoding and memory consolidation.

From the perspective of my broader interests in my project, schemas have been utilized in applied ways in many settings – for example, to aid learning such as languages.

For example, one study paired words with familiar pictures to enhance word learning (Havas et al., 2018). As highlighted in the background sections above, music is made up of tones and temporal intervals that follow strict and regular hierarchies and music has many of the schematic properties of language and semantic concepts, with a distinctive temporal component. Although virtually untested in the broader literature, I in this thesis hypothesized that music that followed predictable structure and hierarchies could therefore provide a temporal schema, that would – like other forms of schemas – confer a benefit to associating new temporal information. This study aimed to test if an association between familiar and regular music provided a sequential structure for encoding sequential associations of non-musical memory.

CHAPTER 3. METHOD

3.1 Participants

Fifty-one participants were recruited from Georgia Institute of Technology volunteer pool. One participant quitted during the task and two participants were excluded due to lack of responses during the task, leaving 48 participants aged 18-24 years (25 females, 23 males). Subjects were pre-screened so that no subject had any hearing problems, abnormal uncorrected vision, or basic music recognizing problems. No subject had learning disability, attention disability, history of other neurological or psychiatric disorder. Participants gave informed consent to the procedures approved by Institutional Review Board before they participated.

3.2 Stimuli

To address my research question, there were two levels of independent variables. The first was the familiarity of music. In the task, a half of the musical pieces (18 in total) were learned on the first day, with the specific compositions pseudo-randomly assigned for each participant (keep distribution of the second level variables even). Here, unfamiliar music could be the reference condition for whether familiarity in music modulates its effect on parallel sequence learning.

The second level of variables was the regularity of the music. To test the effects on visual sequence memory of music in which the tones and temporal intervals were arranged with different levels of regularity and hierarchy, I used three auditory conditions in total for comparison. All music stimuli were composed manually based on following rules. <u>The reference (control) condition had each piece of music/audio (12 in total) been simply composed of monotonic sound</u> – a single note was played 8 times at a steady interval, once per second. This monotonic scale was an appropriate control condition for my research question because it was technically a simple type of music sequence, retaining regularity and predictivity, but lacks the musical hierarchical or harmonic features (Orio & Schwarz, 2001; Wedin, 1972) which I hypothesized above should modulate (benefit or disrupt) parallel encoding of information. To this end, in the experimental conditions the music was an 8-second melody (24 in total) each played by one instrument (flute, clarinet, or piano). Each piece of music contained 14 to 24 notes. Each piece of music had a tempo of 60 beats per minute. Resembling a brief piece of music as might be played in an advertisement paired with slogan, here each piece will be similarly composed – brief, but musical and memorable. <u>In the first regularity condition (regular)</u> half of the music compositions had an organized and predictable structure within a specific key. The root notes of each piece of regular music followed chord progression following western classical music tradition. Tones of the same piece of music were in octave relationship. To test the idea that the regularity of the sequence structure matters for schema learning effects, <u>the second regularity condition (irregular)</u> was made of regular pieces of music composed as above, but then transformed by pseudo-randomizing the sequence the notes and the temporal intervals between notes that occur – that was changing the occurrence of each notes and thus the root notes would not follow a chord progression and the temporal interval between notes were not proportional to each other. However, the keys belonging to one piece of irregular condition music were still in octave relationship and thus were not dissonant. Because of this, the irregular music was not auditory "noise" - it did sound broadly like music and was still memorable, but critically it no longer retained the musical syntax: regular temporal structure and predictability characteristic of the regular/tonal condition.

One potential concern is that I composed the regular music based on western classical music training, and not synthetically using an algorithm. Because of this, I conducted follow-up analyses of statistical properties of my stimuli to verify that these "naturalistic" music stimuli and their irregular condition variants differed in terms of my operationalization of music "regularity" (see Results).

The critical test of my hypotheses came through differences in learning of visual sequences paired with the music stimuli described above. For the arbitrary visual sequences, all the shapes were novel, abstract, and irregular, made up of either lines or curves. There were 36 sequences of 4 shapes – 144 different shapes overall (example shapes in Figure 1).

After the task, a brief questionnaire was given to each participant with questions relating to the amount and form of musical training experience and subjective ratings on how the regular and irregular music affected their emotion and attention during the task.

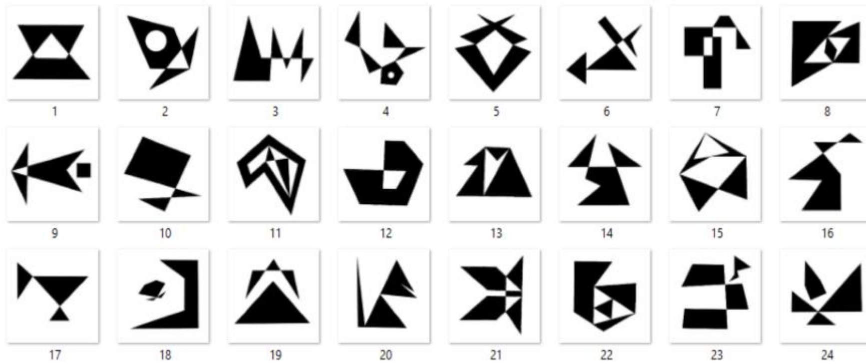

Figure 1 Example visual stimuli used in the experiment

3.3 Experiment Procedures

This was a two-day study containing three parts: music learning, visual encoding, and retrieval. In brief, the participants learned some music (instrumental; no lyrics) during the *music learning* phase. The following day, during the *visual encoding* phase, they learned novel abstract shape sequences paired with music. In the end, during the *retrieval* phase, a test was given on the memory of the visual sequences without any auditory cues.

3.3.1 Music Learning Phase

The main goal for this phase was to implement my manipulation of music familiarity. Here, participants gained a solid memory of some music before visual sequence learning. This stage was accomplished the day before visual stimulus sequence learning so that the participants had a sleep cycle to help consolidate the memory of music. Subjects were asked to memorize 18 novel music stimuli explained above. The music learning task contained 2 parts – on the first part, the subjects listened passively to all 18 music sequences, each in a loop with an unlimited number of repetitions. Their job was to indicate by pressing a key when they felt they had memorized each music composition. Once they finished passively learning all the music, they then proceeded to a music recognition task (procedure shown in Figure 2), which helped them to solidify their knowledge of the composition through retrieval/practice effects.

In this retrieval/practice task (Figure 2), for each trial the subject first heard the first 2 seconds of a piece of previously encountered audio, followed by five consecutive auditory recognition choice questions. The remaining 6-second music fragment was divided into 5 recognition pieces. For each of the 5 blanks, the correct music fragment was presented with a lure choice. The subject needed to re-compose the music by choosing the correct piece for each blank and filling all the five blanks for that music. This test effectively asked the participant to retrieve the music using provided optional musical pieces (an accommodation for the fact that we could not ask them to reproduce the music with, for example, a digital keyboard, particularly since they did not all have prior musical experience).

When the subject made any mistake(s) on the retrieval/practice task, the subject would hear to the correct music again in a loop until they pressed a key indicating that they now knew the piece. The subject then was tested specifically on the previously-incorrect music compositions again in the following experimental run after all 18 compositions had been tested. Across runs, once the subject correctly reproduced a specific music composition on a retrieval task repetition with no mistakes, that music sequence would not be tested again in subsequent runs and was marked as '*learned*'.

As such, this music memory task was designed so that every one of the 18 pieces of music was practiced as many times as needed for a participant to reach an ability to perfectly re-compose them from a forced choice design. The memory task was therefore inherently self-paced. In practice, due to the individual differences in music sensitivity and perceptive ability that I was interested in (see Introduction), some participants did not learn all 18 music compositions to 100% accuracy on these re-composition retrieval/practice questions. Thus, for practical reasons I put a 2-hour threshold so that participants didn't get trapped in the task forever. To facilitate analysis of the subsequent visual sequence learning data, I therefore marked the music sequences that were still re-composed with mistake(s) by the end of the encoding phase as '*unlearned'*. Importantly, failing to achieve perfect re-composition was not equivalent to the music piece being unfamiliar, and so I therefore labelled all 18 pieces of audio present in the encoding phase as 'familiar' sequences for subsequent analysis. In summary, modulated by individual differences in ability, my *music learning* phase resulted in 18 "familiar" pieces of music, with a proportion depending on the individual being labeled as "learned" vs "unlearned".

On the second day, prior to the Visual Encoding phase (see next section) the subject was first given a music memory recognition task for the 18 pieces of audio heard on the first day. This enabled me to identify "familiar learned" pieces that might nevertheless had been forgotten by subsequent day, given finer grained information on when music memory influences other sequence learning (if at all). To test retention of each piece in a reasonable time frame, the subject was given a forced choice between 3 complete 8s audio samples and asked to select the studied music composition from two lure versions of that sequence. The lures will be composed by replacing or shifting a few (1 to 3) notes in temporal dimension from the original music. The subject got tested on each familiar music piece only once. I marked the correct trials as ‘retained’ sequences and incorrect trials as 'forgotten' sequences.

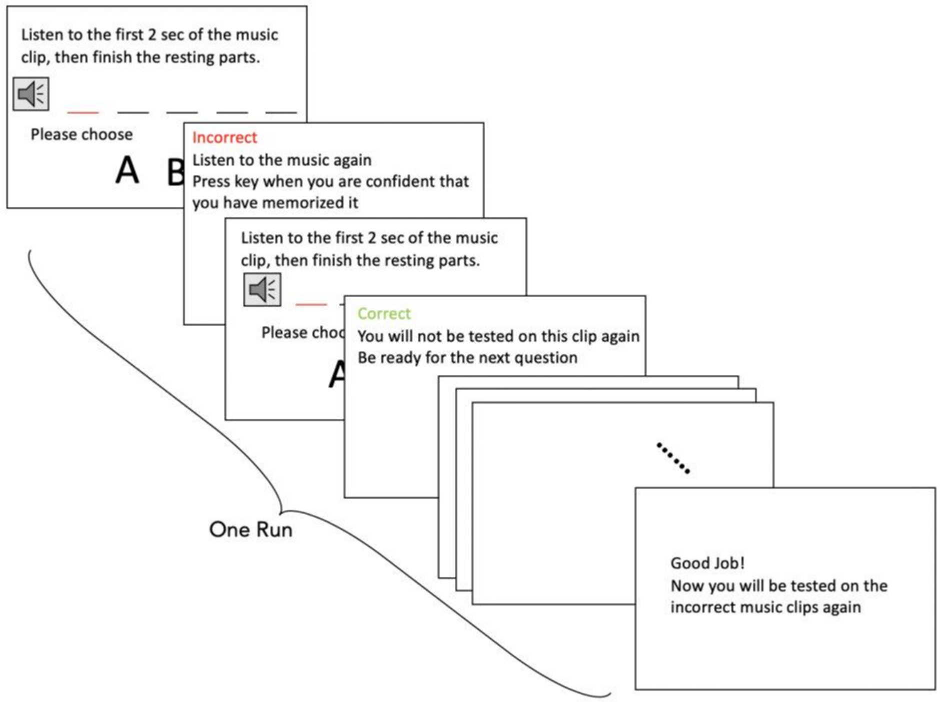

Figure 2 Musical sequence encoding task paradigm used in day 1

3.3.2 Visual Encoding phase

The critical test of my hypotheses came about in the Day 2 Visual Encoding phase. Once participants had been familiarized with the 18 pieces of music (and demonstrate varying degrees of retrieval and retention success as noted above), the main Visual Encoding (learning) phase of the experiment began. Figure 3 showed the paradigm of this learning phase. During this phase, participants learned 36 distinct sequences of four novel abstract and irregular shapes. Each shape sequence was paired with one piece of 8-second music, making each shape correspond to 2 seconds of the music. The subject learned all 36 sequences with randomized order within a run, with 5 runs total (therefore there were 5 learning repetitions per visual sequence). For each trial, the testing screen successively present the 4 shapes in the correct order while the paired music was played. Then the participant saw a blank screen, holding the shapes in their mind, and listened to the music again, but played in a faster tempo of 6 second. Right after this 6s break, the subject would repeat the sequence he/she just learned by sequentially clicking keyboard bottoms that correspond to the shapes. Once each shape/bottom was clicked, the corresponding 2 seconds of music to the selected shape would be played simultaneously. The subject would see feedback of whether he/she repeat correctly or not after submitting the answer. Critical for testing my predictions about music serving as a temporal schema – 18 of the 36 visual sequences were paired with the 18 familiar music sequences, and rest were paired with 18 novel unfamiliar music compositions (with the same general properties). This enabled me to characterize arbitrary visual sequence learning curves in the *Encoding* phase as a function of having familiar/unfamiliar music as a backdrop, as well as whether the music was regular or irregular (scrambled and less predictable).

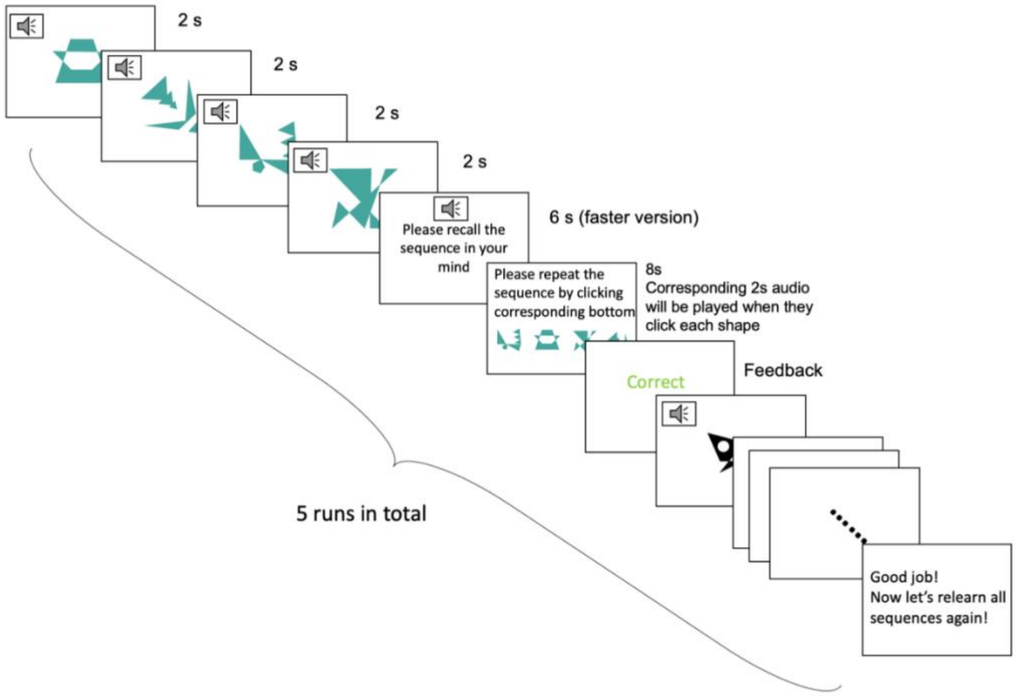

Figure 3 Visual sequence learning task paradigm used in day 2

3.3.3 Retrieval Phase

The retrieval phase tested participants on their final memory of both the new visual sequences and the music sequences at the end of the experiment. In the visual sequence memory retrieval task, each subject was presented visual sequence shapes and needed to reorder the shuffled shapes into the correct temporal order (the hallmark of a sequence). In each trial, the screen showed 4 shapes aligned in a row in shuffled order. Participants had 12 seconds to type in the corrected temporal order number for the shapes displayed. Without receiving feedback, they were allowed to change their answers within the 12s window. There was no music played this time, providing a "pure" test of the final memory for the visual sequence structure. Each sequence was only tested once.

The musical sequence memory task, on the other hand, was a form of error detection task. In each trial, the participant first heard a version of one piece of music (8 seconds). Only regular and irregular music would be tested (24 in total). Because the control condition was monotonic sound lacking tonic harmony with steady temporal intervals, memory for the control condition was tested (in simple terms, there was no opportunity for an error in this condition). There was a potential difference (error) in each second of the regular and irregular music conditions played in this task compared to the original version. The difference could come from changes in both tones and temporal interval. The numbers of error varied across music because the goal was to test if the subjects could not only to detect the error but also to identify correctness: some music had no error while some has up to 8 errors. The music would be played twice, back-to-back, for each test trial – because the participants needed to respond extremely fast once they detected an error, they would hear the music twice so that they could be prepared to respond after first time's detection. During the first repetition, the participants would not be asked to make a response, but simply to pay close attention to detect which second(s) of the piece contain(s) errors. During the second replay, the participants needed to click a button as soon as possible when the error occurred while the music was played. Each piece of music was tested once. The music memory task used a different probe from the re-ordering test for the visual sequence memory task because 1) ordering shuffled 4 parts of music might be easy due to humans comprehension sensitivity to music –for example, the syntax and learned schema for music that humans acquired over the course of lives could make putting a song in order too easy, and 2) this error detection task required faster reaction and can provide more precise measurement for music memory (score of 0 to 8 here).

3.4 Follow-Up Baseline Task

Based on the positive results of familiar and regular music's effect on visual sequences learning (see Results), I further considered how visual sequence learning performance was influenced by listening variety levels of music when compared to not listening to any music at all. To achieve this, I ran a baseline version of the experiment by replacing the control (monotonic) condition with a no-sound condition. I recruited another 20 participants for this follow-up task following the same recruitment procedures (11 females, 9 males, aged 18-27 years). Analyses for this group were identical to the main task.

3.5 Analyses and Hypotheses

The main question asked by the project was whether familiar music and regular music facilitates visual sequence learning compared to unlearned and/or irregular. If so, I would further aim to understand the features of music giving rise to music's enhancing or impairing effects. Based on this goal, I analyzed visual memory retrieval performance and visual sequence encoding performance (learning speed) as a function of the experimenter-created music conditions described above, as well as in a follow-up characterization of the specific regularities and properties unique to each regular and irregular composition. I further assessed the effects of music on visual memory as a function of individual differences in musical skills.

Based on previous findings from the literature, I hypothesized that familiar and regular musical pairing could give rise to an improvement in learning arbitrary sequences of new information (here, abstract visual sequences), relative to novel and irregular music (which were more limited in both the “learned” and “predicably-structured” properties of schemas). To be more specific, I hypothesized in both the final visual memory task and during visual sequence encoding, the participants would learn visual sequences the fastest and with stronger final memory when paired with learned and regular music and the worst when paired with unlearned and irregular music. Because the control condition was not memorable and not dynamically changed in its tones and temporal intervals – thus was matched to both in terms of providing background sound but was at the same time unable to provide cues about temporal position or order - it provided an ideal comparison condition for the different statistical and familiarity properties of my regular and irregular conditions. I specifically hypothesized that familiar and regular condition would show improvement compared to the control condition. In contrast, there was not a clear prediction in the literature for the effects of unfamiliar and irregular conditions compared to the control condition, although they could still show a benefit – due to retaining temporal order information (albeit less predictable), or they could provide distraction, interference, or a division of resources that might even have a negative impact on visual learning. I expected to see a stronger effect overall in people with music training history because they were more sensitive to music and could possibly gain more solid schema for music whereas it was easier for them to perceive the syntax error in irregular music.

To test these hypotheses, I conducted separate analyses on visual sequence retrieval and visual sequence learning curves from their encoding phase. Visual sequential memory retrieval performance was represented by both the averaged retrieval accuracy across trials and averaged reaction times – with faster correct-trial reaction times indicating a stronger memory strength. I predicted that the overall visual memory performance would be good across participants due to repeated practice during the encoding phase, but that the visual sequence retrieval accuracy would be significantly higher in the well-learned music conditions than the unlearned conditions, reflecting a schema benefit. I further predicted higher visual sequence retrieval accuracy in the regular music group compared to both the control and irregular group, due to these compositions comparison conditions lacking the regular temporal structure important for music "schemas" and syntax. Moreover, I hypothesized that these effects to be moderated by the subjects' musical skill level, such that participants with more musical training would be more strongly influenced by the properties of the music.

Given the design of long encoding period, I predicted the potent schema benefits would also be reflected in the learning curves of the encoding phase leading up to the retrieval test. During the encoding phase, because participants retrieved the visual sequences at the end of each trial, I used their performance (either correct or incorrect per each trial) to derive learning rates. To measure the learning speed, I included an additional measurement of learning speed – at what learning stage each individual visual sequence was successfully learned (defined as the participant always correctly re-ordering the sequence in the following runs after that learning stage), called *successful acquisition phase*. For this, I expected to see that participants learned visual sequences paired with well-learned and regular music in the earlier stages of the learning phase than novel and unpredictable irregular music. Because this effect depended on learning a composition-specific "schema" above and beyond learned rules about music from our lives in general, and on processing the music composition's regularity/statistics, I predicted that this effect differed across individuals with different musical training experiences and music.

CHAPTER 4. RESULTS

4.1 Music Familiarity Validation – Final Music Memory Task

First of all, I tested participants' memory for all regular and irregular music at the end of the task in order to verify that participants in general did have better familiarity for old music that was shown on Day1. Figure 4 showed the distribution of task score in each condition. For all pieces of old music, the average score of the error detection task was 6.2 out of 8 and the new pieces had an average score of 4.8 and expected significant difference from old music ($t_{df=1222} = -3.917$, $p < 0.001$), given that these were novel (or still being learned) on Day 2. There was also significant difference between memory performance for *regular* and *irregular* pieces in both new and old condition (new: $t_{df=609.2} = -3.824$, $p<0.001$, old: $t_{df = 609.2} = -2.345$, $p < 0.01$). These were an important result because 1) the result of better score for old music validated the design of training participants to memorize half of novel music stimuli in order to develop two levels of music familiarity. 2) the result showed that old regular music gained better score than old irregular music. It was as I anticipated that regular compositions could be learned more precisely because of their syntactic structure. However, this study aimed to test regularity's influence via a form of schema effect, thus there was a need to equate the visual sequence learning conditions such that they did not differ in background music familiarity but only in the intended perceptual/sequential structure differences of those compositions. I further explained and targeted this problem in the next section. 3) although participants never intentionally learned Day2's music – the only chance they listened to the music was during the visual learning and supposedly their attention was on learning the visual sequences, the participants still gained partial memory for the new music. It again suggested that humans were able to encode music unintentionally.

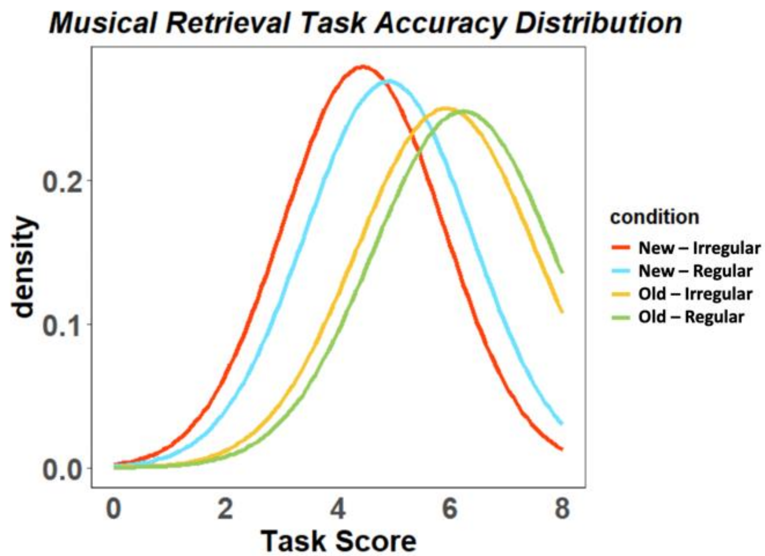

Figure 4 Musical retrieval task accuracy distribution

This is the final music memory test results. The task was an error detection task giving back a score from 0-8 where 8 represented no mistakes and 0 represented zero correctness. The plot shows the accuracy distribution for each music group (excluding the control condition, which cannot have errors/be tested). Y-axis represented the proportion of music pieces distributed at the corresponding accuracy score.

4.2 Music Learning Skill and Music Training History

People vary a lot in their music skills and numerous studies have provided evidence for an association between music training and other cognitive abilities such as visual processing speed and attention (Degé et al., 2011; Rainey & Larsen, 2002; Roden et al., 2014). As a result, individual differences in music skills should be considered in this study, and should be tested particularly in two aspects: 1) how music training affected Day 1 music learning task and resulted in different levels of music memory. 2) in the following analyses of visual memory, how music's effect on parallel task was mediated by music training.

As predicted, subjects' music learning performance on the first day varied considerably. 25 out of 48 subjects finished perfectly learning all 18 pieces of musical sequences (re-composed the music by filling the five blanks perfectly), although the majority of participants learned most of the pieces (minimum: 6 pieces perfectly learned, mean: 15.4 pieces perfectly learned). The numbers of pieces of music learned perfectly on the first day across subjects had a standard deviation of 3.37. This variability again indicated individual differences in levels of music learning skill and provided me with an opportunity to distinguish, at least for some people, visual sequence learning when paired with 'familiar unlearned' vs novel music conditions in the learning phase. To further explain this variability, I looked at the correlation between training history and music learning performance on Day1. Each subject's performance (r) on Day1's re-composition task was defined as the ratio of perfectly learned pieces of music out of 18 minus total time used to finish the whole task z-scored across subjects ($r = n/18 - \text{z-score}(t)$). Each subject's music training history was defined as cumulative years of training in playing instruments or/and singing. I found a significant positive correlation between years of training and music learning performance on Day1 (Pearson's $r = 0.49$, $p < 0.001$), indicating people with less music training spent more time on music learning and ended up retrieving the music with less accuracy. In the following analyses having individual differences in music skill as a factor, I grouped the participants into *Experienced* and *Less Trained* (in music) using 7 years, the median music training length as dividing point.

I also compared the Day 1's learning performance between regular and irregular music. It turned out that subjects in general significantly learned more regular music than irregular music with perfect performance ($F = 12.06$, $p < 0.001$). Because participants learned less irregular music perfectly, in order to compare how schematic regular and irregular music affected visual memory with the factor of regularity being the only difference instead having difference levels of knowledge, I decided to selected music as the schematic music (*learned* music) using a strict criteria – the music, no matter regular or irregular, needed to be perfectly learned on Day 1 so that the subject memorized and were able to retrieve each small details of the piece.

When all Day 1's music was tested again on the second day (excluding monotonic sequences), the accuracy percentage ranged from 33.3% to 100% with a mean accuracy of 85.1%. There was no significant difference between accuracy of regular musical sequence versus irregular musical sequence across subjects (repeated measures t-test: $F=0.816$, $p = 0.367$), indicating that the participants retained memory of most music after a day and the deliberate training on both music types kept these conditions well-matched on the second day and before the visual sequence learning task.

In the following analyses, in summary, I labeled musical sequences as '(well) learned' stimuli if the participants successfully acquired that piece on the first day and successfully recognized it on the second day. Again, strict criterion was used because the hypothesis asked how schema (music here) could affect new memory formation and thus only music with good memory were included in schematic group. I collapsed across the rest of the music pieces from Day 1 and all 18 other sequences that were not present on the first day into 'unlearned' music. All control (monotonic) music appeared on Day 1 were marked as ‘learned’ and ‘unlearned’ if not appeared on Day 1, for better control and comparison. Other than regularity of music, ***learned*** vs. ***unlearned*** music made up the second independent variable, level of music familiarity.

4.3 Visual Memory Retrieval Test

Participants scored either correct or incorrect for each trial during the visual memory retrieval task depending on whether they successfully re-ordered each sequence’s 4 shapes into the correct order. Overall visual task accuracy percentage across subjects ranged from 58.3% to 100% and 90% on average. I expected this high level of performance (on average) because participants practiced the relatively brief sequences during learning for almost two hours. To test whether the visual learning results differed across conditions, I used repeated-measures ANOVA on the memory accuracy affected by level of music familiarity, the regularity condition, and the subject music training. Results showed a significant interaction between music familiarity and regularity ($F = 4.168$, $p = 0.0156$, see all other ANOVA results in Appendix. Table 2 Repeated measures ANOVA results). Pair-wise simple effect comparisons using Tukey’s HSD showed that learned irregular music condition had significantly less visual retrieval accuracy than unlearned irregular/regular condition, plus there was a trend that learned-irregular music had less accuracy than learned-regular music (Figure 5.a, learned irregular < learned regular: $p<0.01$, learned irregular < unlearned regular: $p<0.05$, learned irregular <learned regular : $p =0.0631$), indicating a disruptive effect of learned irregular music on pairing visual sequences learning (see Appendix. Table 3).

Next, I analyzed the reaction time differences among conditions. Many studies have shown that reaction time (RT) in memory retrieval tasks can potentially represent the strength of memory (Gimbel & Brewer, 2011; Raab, 1962; Robinson et al., 1997), especially for correct trials – with a faster reaction or response in successful retrieval implying a better memory and a higher confidence. Thus, I measured the length of reaction time (time taken for participant to place the shapes in temporal sequence) for correctly retrieved trials only (shown in Figure 5.b). Using the same statistical test, I found a significant effect of regularity ($F = 7.828$, $p < 0.001$) and an interaction between regularity and music familiarity ($F = 11.743$, $p < 0.0001$). The remaining ANOVA test results were shown in Appendix. Table 2. Comparing the simple effects between pairs, interestingly, I found that within the *unlearned* condition, the *irregular* conditions had significantly faster reaction times than *regular* condition for correct trials and it was trending that *irregular* condition had faster RT than control condition (*unlearned irregular* RT<unlearned control RT: $p = 0.06$, *unlearned irregular* RT < *unlearned regular* RT: $p <0.001$). For visual sequences encoded with *learned* musical sequences, the *regular* condition had a significantly faster rection time than the control (learned regular RT < learned control RT: $p < 0.05$). Refer to Appendix. Table 4 for all between condition comparison on their reaction time for correctly retrieved sequences. In addition to interactive effect between regularity and familiarity, I also found significant differences in the above ANOVA as a function of music experience (my *experienced* vs *less-trained* factor) ($F = 7.13$, $p < 0.01$). It showed that experienced subjects generally retrieved remembered(correct) visual sequences with less reaction time than less-trained subjects (less trained vs. experienced,

$t_{df=1760.6}$ = -2.66, p<0.01), implying that more music training might be associated with a stronger learning progress in this setting across conditions.

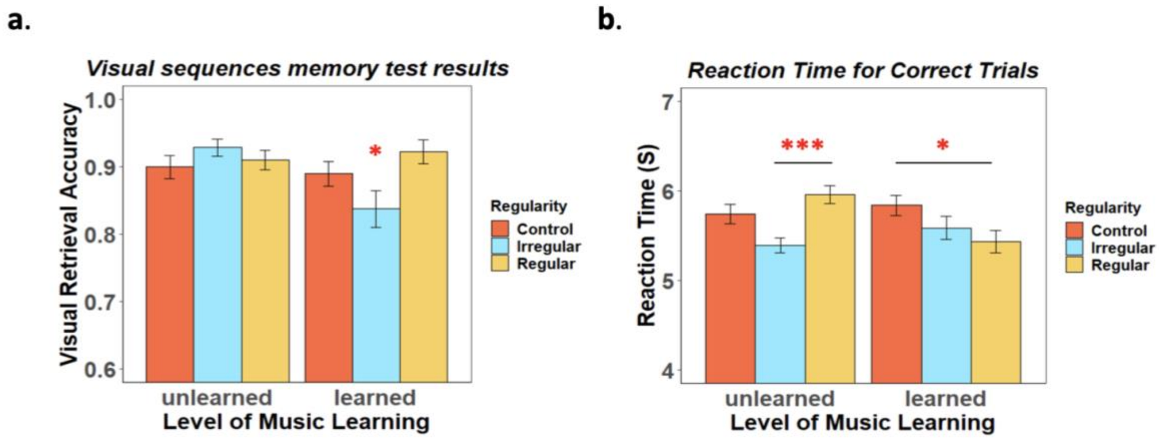

Figure 5 Final visual sequences memory test results. a) Visual sequences memory test results. b) Reaction time for correct trials

a) this bar plot represents average retrieval accuracy for visual sequences for each group comparing learned vs. unlearned, three regularity conditions. The repeated measures ANOVA revealed a significant effect of music familiarity and interactive effect of music familiarity and regularity. The pair-wise comparison indicated the difference came from the learned-irregular group. **b)** this similar bar plot represents average reaction time for correctly retrieved trials only for visual sequences of each group comparing learned vs. unlearned, three regularity conditions. The repeated measures ANOVA showed significant effect of regularity and the interactive effect between it and music familiarity. Pairwise comparison results: unlearned irregular correct trails had quicker reaction time than unlearned control and unlearned regular conditions, whereas learned regular condition had quicker reaction than learned control and unlearned regular condition. **(p<0.001: ***, p<0.01: **, p <0.05: *, 0.05< p<0.1: ·)**

Summary of findings relevant to hypotheses: Together, the average accuracy and RT results indicated two results consistent with my predictions: 1) when visual sequences were encoded with learned music stimuli, their subsequent sequence retrieval accuracy was disrupted by *learned irregular* music at encoding, while reaction times for correct trials (a proxy for sequence memory strength in this task) significantly benefitted from *learned* *regular* music. Unexpected for my theoretical framework, unfamiliar irregular was also associated with improved subsequent visual sequence RT.

4.4 Visual Sequential Learning Curves

Having the participants practice sequence retrieval during each trial of the visual encoding phase enabled investigation of visual sequence learning qualities such as learning speed as a function of the music conditions. Following my proposed analyses, I computed the *successful acquisition phase* (the run number between 0-5 runs) separately for each visual sequence (when the participant for the last time made an error on retrieving that visual sequence). This value represented when the participant had learned each sequence, and the cumulative acquired sequence count for each encoding phase for each condition (e.g., Learned_Irregular) provided insight into visual sequence learning rate. I ran a repeated measures ANOVA on participants' sequence acquisition learning rates to test whether this phase value was affected by the familiarity, regularity, and music experience. Consistent with the average performance results in the prior section, there were significant effects by music familiarity and regularity (music familiarity: $F = 5.181$, $p < 0.05$; regularity: $F=6.244$, $p<0.01$). There was also a highly significant interaction between music familiarity and regularity ($F=9.901$, $p<0.001$). Interestingly, there was no difference between *Experienced* and *Less Trained* groups (Appendix. Table 2). I therefore collapsed across the two music training factor levels and visualized this result by plotting out the learning curve using the cumulative proportion of acquired visual sequences for each condition (Figure 6). To simplify the visualization, I combined *learned control* with *unlearned control* condition performance for each phase because 1) there were no statistically significant differences between them (TukeyHSD: $p = 0.9113$, refer to

Appendix. Table 5) plus 2) conceptually, the monotonic sequences, without dynamics changes in notes and temporal intervals, could not be learned and schematized thus were identical in both familiar and unfamiliar conditions. As shown in the Figure 6 and consistent with the average performance analysis in the preceding section, pairwise comparisons indicated that *learned irregular* music paired shapes sequences were learned the slowest – beginning and maintaining a significant deficit from the other conditions until the 5th encoding phase (see detailed between condition t-test results in Appendix Table 6). Critically for this thesis, it was important to stress that these findings revealed that the disruptive properties of irregular music on parallel sequence learning were driven, on average, by the temporally-irregularly structured music being familiar and well-*learned*, suggesting irregular music is not simply "distracting audio". There were no other statistically significant differences between the rest of the curves (see Appendix. Table 5 and Table 6).

<u>Summary of findings relevant to hypotheses:</u> Together, the learning-curve analysis of correct visual sequencing rates over encoding events generally corroborated the average accuracy results detailed in the prior section. Namely across all subjects when the visual sequences were learned when listening to paired *learned irregular* music, the visual sequences would be encoded the slowest and the worst.

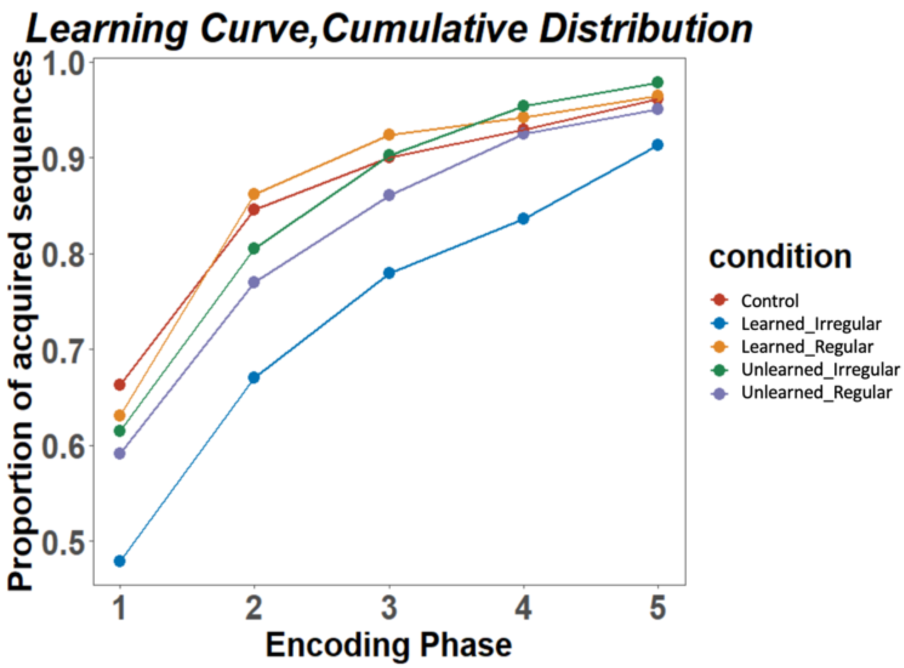

Figure 6 Learning Curve, Cumulative Distribution

This figure represented the cumulative learning curve during visual encoding. The plot visualized the average proportion of visual sequences already learned so far at each stage as a function of each music condition (e.g., the very left bottom dot – participants on average learned 48% of the visual sequences when paired with music from the *learned irregular* condition after the first phase 0. The repeated measures ANOVA test revealed the significantly slowest learning happened in the learned-irregular condition.

4.5 Clustering Analysis on Visual Learning Behaviors

Interestingly, contrary to my predictions, the group-level analyses above emphasized effects on visual sequence learning for the *learned irregular* condition (with benefits of the learned regular condition manifesting in superior RT for (otherwise comparable to control condition) correct sequencing test trials). One possibility was that the average effects encompassed multiple meaningful clusters of distinct individual behavior/learning profiles. Indeed, in the visual learning task I predicted divergent learning behaviors depending on music experience, and although I did not observe evidence of this at a broad level, I considered the possibility that this could still be a factor for the profile of how someone's visual learning responds to the different music conditions.

To test this idea, I ran a clustering analysis by using a K-means nearest neighbors algorithm, and the 'elbow method' (Kodinariya & Makwana, 2013) to select the most meaningful number of clusters. The features of each sample used were participant's average visual sequence accuracies for all conditions (learn regular, learned irregular, unlearned regular, unlearned irregular, control) during the visual sequence encoding phases. Because five participants finished the task on Day 1 with 0 perfectly recomposed irregular music, they lacked features for the *learned irregular* condition and thus the algorithm automatically dropped these five samples, leaving 43 samples in this analysis.

The KNN analysis uncovered 3 clusters of the learning profiles during visual sequence encoding. To visualize the similarity between clusters, I ran PCA (principal components analysis), a dimension reduction analysis and plotted out the clusters according to the first and the second component representing the data, which explained 42.75% variability of samples (Figure 7.a). This plot revealed that one participant learning profile group (cluster 2) represented a more divergent behavior from the other two types of learning participant (cluster 1 and 3 shared more similarity). The important thing for this thesis was what behavioral profiles in my task these different clusters represented. I plotted out each cluster's average learning curve. The divergent cluster was quite small (#2, n=3, Figure 7.c) and therefore difficult to interpret or assess with inferential statistics. Focusing on the larger-sample Cluster #1 and #3, broad similarities between the learning profiles were marked be several important differences. In cluster #1 (n = 18; Figure 7.b), differences from music conditions were restricted to the early stage of learning (run 1) and participants from this cluster showed a trend that they learned the visual sequences paired with learned-regular the best and those paired with unlearned-regular the worst (*learned regular* > *unlearned regular*, $p = 0.025$, all t-test comparisons shown in Appendix. Table 7). In cluster #3 ($n = 22$) on the other hand (Figure 7.d), the most clear differences were between *learned regular* and *learned irregular* (resembling the group average results in Figure 6) – participants in this cluster learned the visual sequences with *learned irregular* music the slowest and with the lowest accuracy across run1-4 (refer to Appendix. Table 8 for all comparisons results), while there was trending that they learned sequences with *learned regular* music better than the other group especially in the middle stage of learning (run 2-3) (Appendix. Table 8).

<u>Summary of findings relevant to hypotheses: </u>participants in cluster 1 learned visual sequences differently only for the regular condition, with opposite patterns relative to the other conditions (i.e., control and irregular) for learned vs. unlearned regular music. Cluster 3 participants showed the predicted effect for my thesis: an opposite pattern for regular vs. irregular condition that was driven by the structure of each condition being learned.

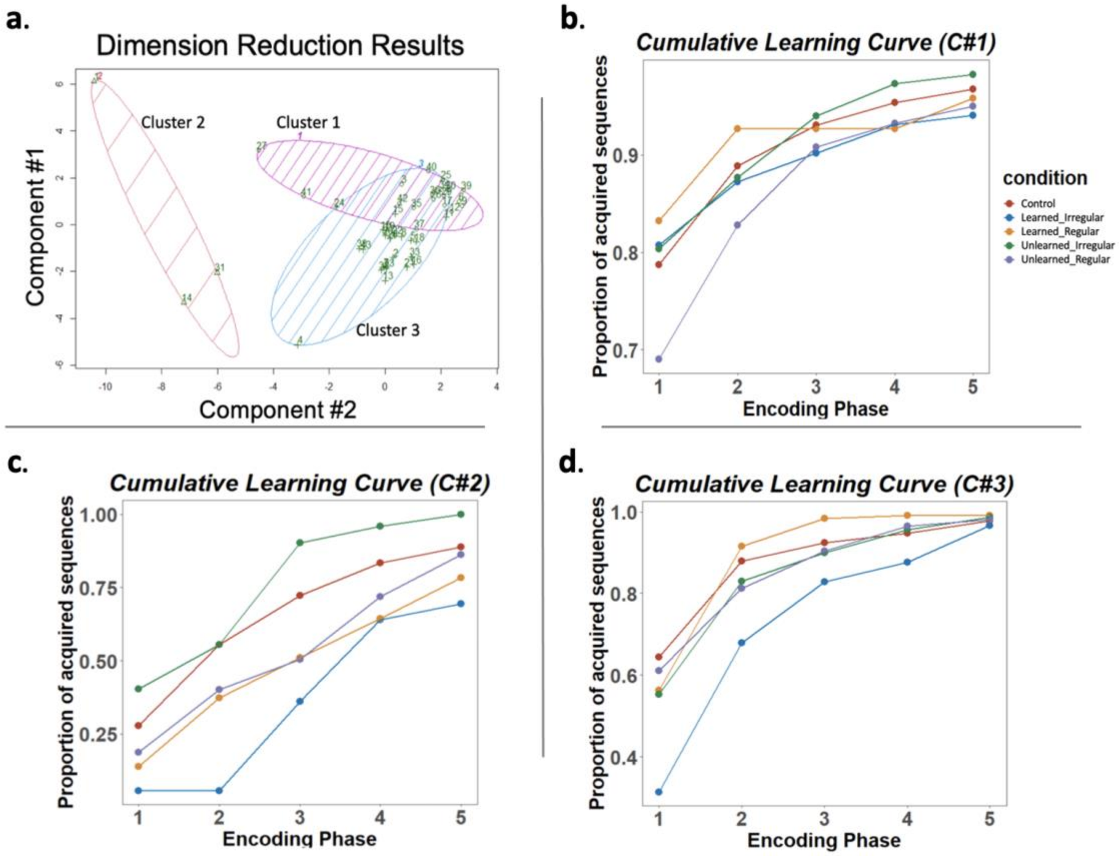

Figure 7 Clustering analysis results: a) Dimension reduction results. b) Cumulative learning curve (Cluster 1). c) Cumulative learning curve (Cluster 2) d) Cumulative learning curve (Cluster 3)

a) after running PCA on features, two most important components were found and this dot plot represent these two components value of each sample. **b-d)** Learning curve for averaged cumulative proportion of acquired sequences at each stage during visual sequence learning for each cluster. Sample size of each cluster were 18 ,3, 22.

4.6 Applying Clusters From Encoding Phase to Retrieval Results

In addition to comparing clusters on their averaged learning curves, I also wondered whether clusters associated with different memory retrieval performance, especially between cluster 1 and cluster 3. Answering this question could be insightful for whether a specific learning pattern in during encoding correlated to a better or worse final sequential memory performance. Figure 8 showed separate plots of visual retrieval accuracy and reaction time for correctly retrieved sequences for subjects from cluster 1 and cluster 3. First of all, I found that across conditions, cluster 1 subjects achieved visual sequence retrieval with better accuracy than cluster 3 subjects ($t_{df=1431.9} = 2.994$, $p < 0.01$). When focusing on reaction time for correct trials, I found cluster 1 subjects generally retrieved these trials with shorter reaction time across conditions ($t_{df=1204} = -4.4179$, $p<0.001$). Both results suggested that cluster 1 subjects gained better visual sequence memory than those from cluster 3. Secondly, I ran pair-wise comparison between conditions regarding to their visual sequences retrieval accuracy and RT for correct trials. Due to small sample size within clusters, I utilized unadjusted pair-wise t-test to reveal the trending patterns in clusters differences that could be explored further in future work. The significant results for all between-condition comparisons were shown in Appendix Table 9. All significant or trending differences were highlighted in Figure 8. When comparing between clusters as well as referencing back to Figure 5 (my whole-sample result), I found similar patterns in both clusters – *learned irregular* condition showed the worst retrieval accuracy while *learned regular* condition showed a tendency for improved accuracy (it is difficult to interpret significance levels due to small sample size). RT for correct trials plot also showed similar patterns to overall dataset – *unlearned irregular* music condition had shorter response times than control, implying a stronger memory on correct trials than control.

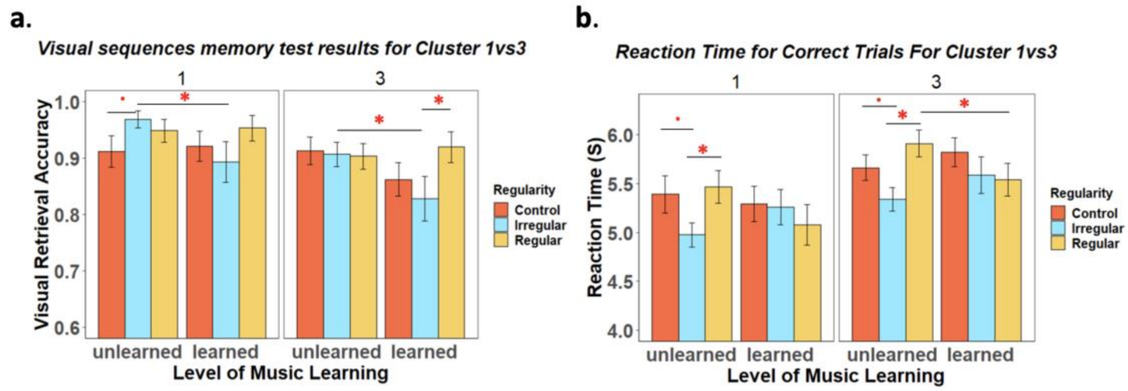

Figure 8 Comparing retrieval results between Cluster 1 and Cluster 3: a) Visual sequences memory test results for Cluster 1vs3. b) Reaction time for correct trials for Cluster 1 vs 3.

(**a**) visual sequences retrieval accuracy: this bar plot represents average retrieval accuracy for visual sequences for each group comparing learned vs. unlearned, three regularity conditions for cluster 1 vs cluster 3. (**b**) reaction time for correctly retrieved sequences: this bar plot represents average reaction time for correctly retrieved trials only for visual sequences of each group comparing learned vs. unlearned, three regularity conditions, comparing cluster 1 and cluster 3. Pairwise t-test was conducted on all possible pairs conditions and on both retrieval accuracy and RT, the detailed results were presented in Appendix. Table 9. All significant and trending differences were marked in the plots using asterisks. **($p<0.001$: ***, $p<0.01$: **, $p<0.05$: *, $0.05< p<0.1$: ·)**

4.7 Why did people show variant learning behaviors?

The clustering analysis revealed three distinct patterns of visual learning that were affected by my paired music manipulation. The critical question was “why?” Based on schema theory, I had predicted that music familiarity (having a memory for the music composition’s structure) and the qualities of the music composition’s structure (its regularity and predictability) would both be factors the influence “schema-like” effects of music on parallel visual learning. The visual learning analysis showed that for most of the participants (cluster 1 and 3) there was a trending improving effect from *learned regular* music on visual encoding - corroborating a benefit of music familiarity and regularity (seen at least the group level with indicators of stronger correct sequence memory; Figure 5.b).

However, comparing cluster 1 and cluster 3 revealed an important distinction - for some participants learning was impaired by a lack of familiarity (cluster 1: the *unlearned regular* music was learned the worst) while other participants appeared more vulnerable to familiar but irregular music. In order to investigate the reasons behind this, I compared key participant variables for these two cluster distributions. In particular, I revisited my theoretical interest in how music training may influence participant sensitivity to music structure and familiarity effects on learning.

Whereas there was absolutely no evidence that gender affected the way participants learned the visual sequences under different music conditions (identical cluster distribution between females and males (Figure 8)), there was a clear effect of different levels of music training. As shown in Figure 8, the more instruments the subjects had learned, the less likely they exhibited a cluster 3 learning profile. I ran a Pearson's chi-squared test with a Yate's continuity correction to establish significance and tested whether people with different music training history were distributed equally across cluster 1 and 3. Subjects with experience with 0-1 instrument (combined for statistical power considerations) exhibited significantly different cluster membership than those who had learned more than one instrument (chi-squared = 3.90, df=1, $p = 0.049$). Though the figure showed that musically experienced group showed less cluster 3 behaviors (Figure 9), the years of experience as a musician did not exhibit the same statistical effect (chi-squared = 0.94, df=1, $p = 0.33$), noting this could due to small sample size but also this median-split approach would be expected to be less sensitive than the other measure since the "less-trained" group could still have up to 7 years of music training.

<u>Summary of findings relevant to hypotheses: this individual difference analysis reveals that music training matters -</u> more training with music was associated with a greater sensitivity to whether regular music in the background was learned or unlearned, and a lower likelihood that learned irregular music would be disruptive to parallel learning. One speculation in relation to my theoretical framework was that the subject who gained more music training would be more efficient at processing the music so that the retrieval of what tones, plus the paired shape, was coming next in learned music (even if irregular) was more automatic. I examined these ideas further in the next section, which formally quantified the structural differences between regular and irregular music.

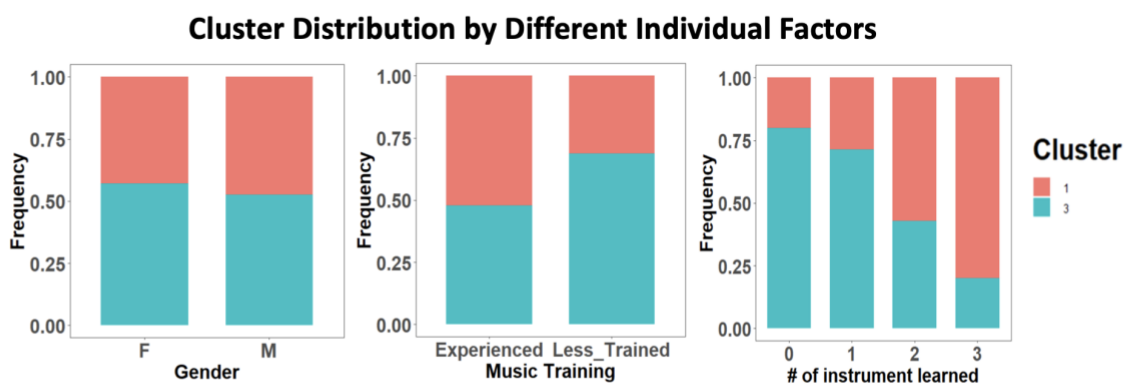

Figure 9 Cluster distribution by different individual factors

This figure compared Cluster 1 vs. Cluster distribution in different groups. **Left)** cluster distribution on subjects grouped by number of instruments learned. 2 subjects who learned 4 instruments were not shown due to small sample size. It showed that more proportion of cluster 3 behaviors was associated with less music training. **Middle)** cluster distribution on subjects grouped by years of music training, 7 years as the dividing point. No statistical difference was found (chi-squared = 0.94, df = 1, p = 0.33) **Right)** cluster 1 and 3 were equally distributed among females and males (chi-squared = 0, df = 1, p = 1).

4.8 What is different between regular and irregular music?

In the previous analyses, the noteworthy observation was the opposing effects of *learned regular* and *learned irregular* music on visual sequence encoding– I observed an improving effect from *learned regular* music (RT analysis in Figure 5.b and visual retrieval accuracy among partial subjects in Figure 8.a) but hindering effect from *unlearned irregular* music (shown across analyses in Figure 5 and Figure 6). Both conditions being well-learned used a strict criterion (Method) so that the music from both *learned* conditions were of high level of familiarity - however, they had divergent influences on parallel visual memory encoding. This result indicated the importance of regularity in determining music's effects. Moreover, my individual differences analysis revealed that music training experience influenced the effect of my music conditions. However, as noted in the Methods, I composed my so-termed "regular" music stimuli from the standpoint of a western classical musician, and not synthetically based on some specific parameters provided to an algorithm, and I generated my irregular music by scrambling the temporal order. Because of this, it was important to verify whether key properties related to tonality and hierarchical structure did differ between my conditions. In this section, I computed statistical properties of the music and validated the intended structural differences of the regular and irregular stimuli used in this study using quantitative means.

To quantify the structural regularity of music stimuli excluding monotonic condition, I used the MIRtoolbox (Music Information Retrieval) from MATLAB (Lartillot & Toiviainen, 2007). Many music studies have tested this toolbox in its ability to distinguish and analyze instrumental regular and irregular music, with these studies highlighting the *key clarity* and *pulse clarity* functions (quantifying the clarity and distinctness of the main key and pulse pattern of the music) as critical properties for regularity (Lartillot et al., 2008; Mencke et al., 2019). For example, Figure 10 represented the envelopes and waveforms of one regular music composition selected from my stimuli and the irregular music transferred from that regular music. The regular music had a clearer pulse pattern because the onset of each note was almost equally distanced with each other while the irregular music showed more various and scrambled tonal distances. Also, in the regular music waveform, a clear accent (larger amplitude) happened about every 1.5s following a few "lighter" notes with similar patterns. This example demonstrated a regular pulse sequence. This figure could not reveal differences between key clarity because in this task, the music stimuli were composed in pairs – each pair of regular and irregular music contained the same notes. As a result in Figure 10 both regular music and irregular shared similar amplitudes of notes. I used MIRtoolbox and computed the key clarity and pulse clarity of all music (excluding control) and ran pair-wise comparison. As shown in Figure 11, I demonstrated that the pulse clarity in regular music was significantly greater than in irregular music ($t(11) = -2.1$, $p<0.05$), on average across pairs. There were no significant difference between the key clarity of regular and irregular music ($t = -1.4$, $p = 0.095$), which was less surprising because both music compositions shared the same notes and even in the irregular music notes were still correlated with each other across time and belonged to the same key. As a result, the greater difference between the regular and irregular stimuli laid in the regularity of the order of the tones and the associated quantifiable clarity of the temporal interval sequence. In language the analog was the words of a sentence being put in an order that didn't follow a syntax and thus would not make sense – each word was clear and comprehensible, but a meaningful structural relationship between them was lacking for higher-level linguistic meaning.

Above, my behavioral data verified the irregular music was memorable, while at the same time this analysis verified it successfully violated the syntax of ordering of tones, and was thus degraded in the hierarchical, predictable features of music.

Summary of findings relevant to hypotheses: Combining it with the previous results, this analysis supported the idea that manipulating the hierarchical feature of the music through pulse clarity was a determinant for the main effects and music-experience-related outcomes on how listening to music altered sequence memory formation.

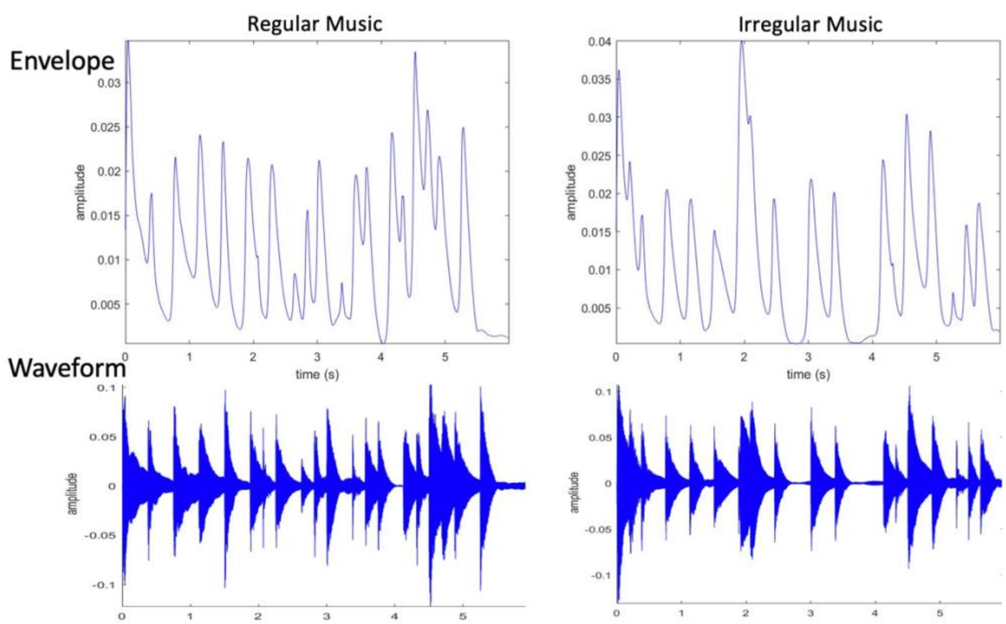

Figure 10 Music envelope and waveform for example regular music stimulus and the paired irregular music stimulus

6s version of the stimuli was used to extract the waveform. Note that the tempo did not change either plots because they showed the amplitude changes over time (simply stretch the above plots horizontally to get 8s version plots of the envelopes and waveforms for the music)

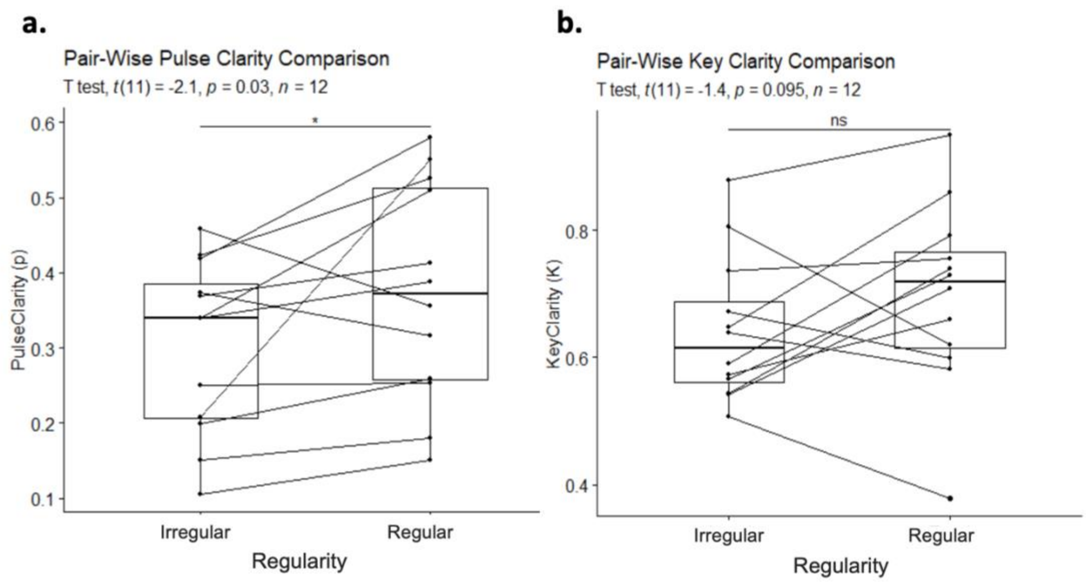

Figure 11 Pair Wise Comparison on Regularity Features: a) Pulse clarity b) Key clarity

pulse clarity and key clarity were measured using MIRtoolbox from MATLAB on both regular and irregular stimuli. Pairwise comparison (t test) was conducted to compare between regular and irregular music in these two features. This figure showed error bar plots of a) pulse clarity and b) key clarity of each regularity condition and each point represented one music. Each pair of regular and irregular music were connected using lines. T-test showed a significant higher pulse clarity in regular music but no difference in clarity.

4.9 Follow-up Baseline Task

The current results showed that *familiar/learned regular* music could result in stronger visual sequence memory encoding, and a *familiar/learned irregular* music could clearly interfere, when compared to the monotonic control condition. One question was how the monotonic control fits into the mechanistic framework I laid out. The monotonic control was the simplest version of musical input: it had no hierarchical structural dynamics that could be harmonic nor memorable – and could thus not be "familiar" either, from the perspective that music schema effects are dependent on the existence of the regularity and familiarity. However, the monotonic control condition did provide a steady one-dimensional metric, akin to a metronome. Although I provided important evidence overall (RT analysis) and some hints in terms of individual differences (music experience effects on learning clusters) for regular benefits relative to this control, one obvious follow-up question was whether the monotonic condition was identical to a no-sound condition as a learning baseline. To address this, I conducted a follow-up experiment with a no-sound baseline task.

In this new participant sample (N=20), the visual retrieval accuracy per condition showed similar pattern to the previous data (Figure 12.a vs. Figure 5.a). Because no-sound control condition could not be learned or unlearned, for each participant, I randomly assigned half of the control condition into ‘learned’ and another half into ‘unlearned’ for easier comparison. This baseline study clearly suggested that the new no-sound control condition did not show better memory retrieval performance than *learned regular* condition. When comparing the reaction time for correct trials (Figure 12.b), there was a marginal familiarity*regularity interaction and corresponding trend (also mirroring the results in the main study) that in *learned regular* condition participants retrieved the sequences correctly with faster RT compared to a no-sound control condition ($t = -1.876$, $p = 0.0602$ unadjusted).

<u>Summary of findings relevant to hypotheses:</u> Overall, despite the smaller sample size these results were consistent with the pattern in the main data, and importantly suggest that the benefits/detriments of regular/irregular relative to baseline observed in my study were not in fact both collectively worse than no-sound (as one might expect if all music conditions were is simply distracting relative to silence).

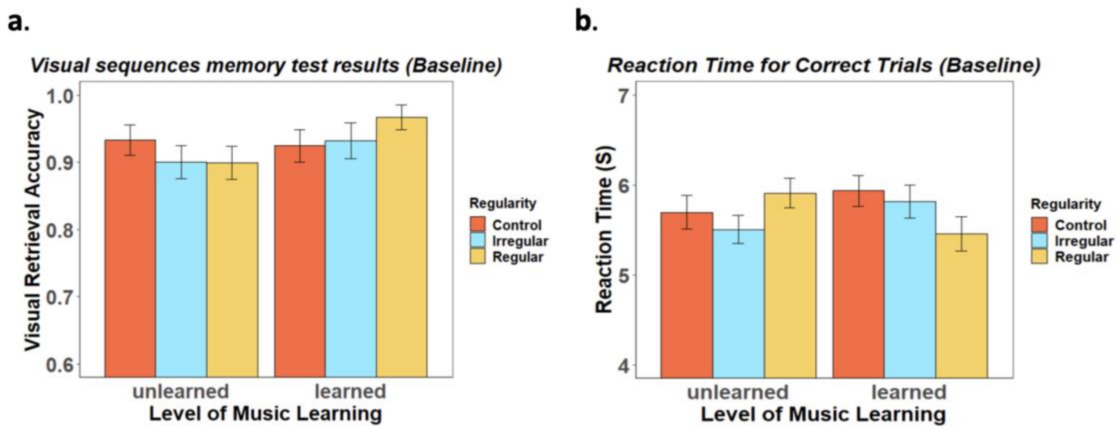

Figure 12 Follow-up baseline task results: a) Visual sequences retrieval results. b) Reaction time for correct trials.

a) this bar plot showed averaged visual retrieval accuracy per conditions with error bar. No statistical differences between conditions were found. **b)** this bar plot showed averaged reaction time for successfully retrieved sequences per conditions with error bar. Learned regular condition (yellow on the right) showed marginally faster reaction time than learned control condition (blue on the right).

CHAPTER 5. DISCUSSION

In broad terms, this study aimed to understand how prior memory and new learning interacted, even across different modalities. Based on the evidence in the literature showing that some music memory was strongly schematized, and that schemas could improve learning efficiency of associated new information (Leman, 2012; Van Kesteren et al., 2012), the study specifically tested how listening to music with different level of regularity and familiarity, which were theoretically important features for music serving as "schemas" and providing stable sequential templates, during visual sequences learning would affect learning results. The results of the study provided insights to applied researches of utilizing music as aids to non-auditory cognitive function such as memory. A concrete real-world example, targeted by my study, was whether listening to our favorite music while studying associative materials from school influenced the learning outcome.

My results suggested yes – and that "familiarity" and "regularity" of the music were both factors deciding the types of influence music had on learning visual associative information. By looking at the final visual sequence retrieval accuracy, I found participants learned the shape sequences with the least accuracy when listening to well-*learned irregular* music compared to other conditions, indicating a significant hindering effect from this type of music. In addition, the reaction time analysis of the final retrieval test revealed two opposite memory outcomes – for the shape sequences successfully learned with familiar music, participants retrieved them faster if the sequences had been learned with regular music. Yet for shape sequences learned with unfamiliar music, participants retrieved them fastest if they had been learned with irregular music. The observations from the visual retrieval task were partly consistent with my hypotheses: that pairing with familiar and regular music was beneficial to learning the associations and the order of new information (novel visual shapes). Partially different from my hypothesis that music conditions with no pre-existing knowledge or structural regularity should show the worst visual sequence learning: the results actually showed that irregular music was clearly detrimental specifically when it was well-*learned*.

Similar patterns were found from the learning curve results during visual sequence encoding. During the encoding period, I focused on the changes of participants' retrieval accuracy of the shape sequences per trial across time/run, which revealed their progress of long-term memory encoding. Across subjects, I observed shapes sequences were learned slowest and with least accuracy when paired with *learned irregular* music, consistent with the final retrieval task results. A clustering algorithm revealed there were distinct patterns of these learning curves - I observed three clusters of learning behaviors affected by different types of music (although one was very rare and thus not amenable to inferential statistics). Comparing the learning patterns of the two dominant clusters, I found that most of the participants showed a trend that they performed the best when listening to learned regular music during visual sequences learning, either during the early stage or the later stages depending on the cluster. This again suggested that familiar and regular music might aid parallel visual sequence memory formation. Conversely, one group of them showed degraded visual sequence learning while listening to *unlearned regular* music, while the other group was negatively affected by the *learned irregular* music (consistent with the overall sample average). Statistical comparisons between musical conditions for the two clusters (Appendix. Table 7and Table 8) suggested that one cluster of subjects showed more statistically-different learning trajectories affected by different levels of music (cluster 3) while the other cluster showed more similar learning progress across conditions (cluster 1). My results of comparing cluster distributions across subjects revealed that music training (how many instruments the participant learned) was an important factor for these clusters in learning differences when music is in the background.

When I used the clustering results from encoding to select and divide subjects into 2 groups, and looked back at their retrieval results, I found similar patterns of these subgroups in both their retrieval accuracy and RT results compared to the overall main dataset. Interestingly, I found that subjects with cluster 1 encoding behaviors retrieved visual sequences with a higher accuracy and faster reaction time (for correct trials) than the other group of subjects across conditions, while the visual sequence performance difference between music conditions remained similar for the two clusters.

Collectively, my results showed an interactive effect of familiarity and regularity – 1) familiar music impaired new visual sequences (association) learning if it was irregular (reflected by poor retrieval accuracy and slow learning progress during encoding), and improved new sequential learning if it was regular (supported by marginally improved accuracy and stronger memory of learned trials indicated by reaction time). 2) Unfamiliar music was a surprising outcome and it showed that unfamiliar irregular music actually benefited visual sequences learning (reflected by RT analyses and partially by stronger retrieval accuracy among cluster 3 participants). My study was the first of this kind, to the best of my knowledge, so more work was needed – but below I considered why.

First of all, the group-level *learned regular* benefits (trending accuracy and significant correct RT benefit) suggested that in order to be schematic and informative to parallel visual learning, the music needed to have both regularity and familiarity. This was also suggested by my individual differences clustering analysis, where the majority of participants across the two dominant clusters were disrupted by music lacked either regularity or familiarity, and marginally favored familiar regular music, regardless of music training. One combined benefit of regularity and familiarity was predictability – these songs had a known and more easily understood structure (which I discussed further below). Secondly, the diverse learning profiles I found between clusters implied that lack of one feature, either familiarity or regularity (e.g., with familiarity but lack of regularity) in the music might distract people from learning the visual sequences, depending on a subject's music training experiences – subjects with more music training were more likely to learn visual sequences the worst if the music was regularity-structured but lacking familiarity, while subjects with less music training were more impacted by the (irr)regularity dimension of music. Thirdly, the surprising outcome was that in the final retrieval test, memory for sequences learned with unfamiliar music was actually strongest (fastest correct RT) if the unlearned music was irregular (lacking both familiarity and regularity). The first two classes of results summarized above have strong consistency with schema theory – that new temporal associations might be better acquired if they could be interwoven with prior and well-structured sequence structures. But this latter outcome highlighted at least one result that was likely due to a different learning-enhancement mechanism. The general pattern of the above results are summarized in Table 1.

Table 1 Results summary of music's effects on visual sequences encoding performances affected by the two music features.

Music's Effect	**Familiar**	**Unfamiliar**
Regular	Beneficial	Distractive
Irregular	Distractive	Beneficial

Some important insights into the underlying factors for why music influenced parallel visual sequence learning in the manner that it did came about when I asked what subject features contributed to the split of clusters. The fact that the visual learning clusters diverged according to music training was striking – participants with less music training tended to be the group driving the main effect of unlearned irregular music harming parallel sequence encoding (cluster 3). By contrast, while more exhaustively musically trained participants also showed a qualitative benefit from familiar regular music, their visual sequence learning suffered from that same regular music being unlearned (cluster 1). Both outcomes could highlight how individual training/experience interacted with the properties of the stimuli themselves. Relating back to what has been suggested in "schema theory", old-memory-congruent information could be learned better and faster (Van Kesteren et al., 2012, 2018). This could explain the latter observation which learned and unlearned regular music had opposite effects. Conceivably, because the music was unfamiliar, it failed to provide a schematic template for benefiting learning new sequence information. But there wasn't just no benefit – there was a detriment. The fact that this was skewed towards more musically-proficient individuals could reflect partial attention being taken from visual information encoding by the coherent "syntax" of the unfamiliar regular music. Musically-trained individuals might be more drawn towards processing the novel regular music than less-trained individuals (a bit akin to trying to study something while someone is speaking a sentence to you in your language). Some evidence from prior studies was consistent with this view, in which participants listened to background music during work, and the results suggested listening to some music could distract attention and caused worse cognitive performance (Chou, 2010; Roden et al., 2014). However, that framework alone doesn't explain the observations in cluster 3, where it was instead the case that, at least in less-musically-trained people, the visual sequences were encoded significantly worse with *learned irregular* music than the other conditions. Both familiar regular and familiar irregular music were memorized equally well within-subject (as shown in the Day 1 and 2 music memory tests) and thus both were able to provide temporal sequence schema to the new memory, and therefore unequal learning could not account for this effect.

One way I sought to understand why *learned regular* and *learned irregular* music showed opposite effects (overall, but particularly in the less-musically-trained subset identified by their distinct learning profile in the clustering analysis) was to run a structural analysis on the music stimuli and I tested what was indeed different between these two types of music. The result showed each pair of regular and irregular music differed in their pulse clarity – this verified that irregular music, although it shared the same notes from its regular music counterpart, lacked a regular temporal structural of the notes. One reason this was relevant was that prior music studies have shown that humans not only have a preference for specific sequences of temporal intervals between notes and regular rhythm patterns, but this had also been shown to be necessary for music's syntax - the hierarchy of high level relationships between music elements which make music "musical" and consistent with its language (Krumhansl, 2000; Krumhansl & Shepard, 1979). Prior literature has emphasized the lack of hierarchy as being characteristic of irregular music compared to regular music (Butler, 1989; Dibben, 1994). An analog to language was that when an English learner tried to read a sonnet, he/she might be able to guess the contents of the sentence by understanding each word, however it might be hard to extract the exact meaning because the unfamiliar grammar failed to provide a clear hierarchical relationship between the words. It led me to speculate that – by being less predictable and interpretable - the lack of hierarchy in irregular music which led to violations of music syntax might affect parallel sequence learning. Studies suggested that neural signatures of processing syntactic relations in both language and music shared similarity, and violations of syntax in language were known to not only fail to produce semantic meaning but also result in neural "prediction error" related indicators in signals like the N400 and MNN (Hagoort, 2003; Koelsch et al., 2013; Patel et al., 1998; Pulvermüller & Assadollahi, 2007). Based on these prior findings, one possible explanation of the hindering effect of the learned irregular music was that with each violation of normal musical syntax, the irregular music elicited a prediction error that drew attentional and mnemonic resources towards the "oddity" in the composition, as the brain attempted to process the conflict between known the structure and what should flow from "correct" music syntax. Cognitively, the conflict in memory of temporal sequences between what was learned (the music) and what should be happening (proper music syntax) might have interfered with the participants' ability of the memory system to encode some other sequence structure in parallel. And the reason why less-trained musicians were affected more by *learned irregular* music might relate to their ability of processing complex musical sequence structures. Similar to the same language analog that people who were trained in ancient literatures and language would find it easier to understand sonnets, although all humans appear to be generally sensitive to music syntax according to the broader literature, training might enable someone with more skills to be able to bridge the tonal and timing structures of a less-regular music and thus process the above mentioned conflicts with less efforts. In this way, *learned irregular* music in my task might become more of a 'neutral' condition in participants with more music training (cluster 1). Still, it was difficult to use the current behavioral study setting to test these possibilities, and therefore I presented them as hypotheses born out of my data and to be tested further with future works. Thus, this Master's thesis set up interesting future directions for my research program: in the next step, I would utilize neuroimaging methods such as fMRI to test the neural mechanisms behind retrieving the music of each condition during parallel visual sequence learning and the respective costs and benefits of those conditions.

In summary, in this thesis I raised a theory that using background music during episodic memory-like learning (here, abstract novel shape sequences) could modulate the new sequence memory strength and learning speed via the psychological schema framework - where the temporal structure from old memories was able to provide a template for better binding new, paired temporal associations. Several key results supported my hypothesis that in order to give facilitating effects to parallel visual sequential learning, the music needed to be familiar and regular/predictive. The results provided rare evidence for schema theory of memory acting cross-modally – such that auditory music sequences influence visual object sequences. This finding can be used to help motivate more future studies examining how schemas can act across other modalities as well (e.g. how motion sequences in dance can provide a schema that modulates learning of other sequences (e.g., visual, as I tested here), or of true episodic memory or even semantic memories). Moreover, speaking to my specific experiment, because sequential memory is an important component of episodic memory – in which humans typically combine separate modalities of information (sights, sounds, emotions) in sequence using a memory system optimized for forming huge networks and associations between events (Eichenbaum, 2013; Eichenbaum & Fortin, 2003; Kesner et al., 2002) – one prediction from my visual sequence study is that other elements of episodic memories may be influenced similarly by music varying the properties that I manipulated. This study, using a simple design (targeting simple shape sequences) provided insight into an applied research question: whether there is a possibility of using music to help people encode the temporal associations components of episodic memory. Future studies could consider utilizing similar task structure and test whether music could modulate more complex associative memory, such as true episodic sequences or spatial navigational sequence learning.

My results also uncovered several surprising complexities in how music schemas affect new learning and partially answered the important question raised in the *Introduction.* about how different types of real-world music might have modulate memory differently in prior studies. In particular, this study showed an interactive effect of music familiarity and regularity on visual sequence memory encoding, where music could be beneficial when the music paired with the new learning featured both familiarity and regularity (as predicted) but also if it exhibited neither of these traits (not as predicted). Moreover, in scenarios where music only lacked familiarity (unlearned regular) or only lacked regularity (learned irregular), the music could disrupt new sequence learning. Which of these two scenarios was disruptive to learning depended on the participants' music training. As such, this study provides novel insight to the question of whether background music (a common feature in our lives!) influences learning. Prior literature on this question has been inconsistent, and most of existing evidence only suggested that music was helpful during retrieval as contextual cues. The current study used a unique design where music was played only during encoding but not during recall and thus provided insight into features of participants and the music itself that could underly that variability in the literature. My results implied that not all types of music were beneficial to parallel learning, with music familiarity and regularity emphasized as key features that influence how building associations between music and sequence memory affects the learning process. Moreover, on the individual level, the study suggested that the pattern in how different types of music influenced sequence learning was shaped by whether subjects had more or less music training. Such perspective might also help provide an explanation for past literature showing different results when testing music's effect on various types of memory – the individual music listener's traits might matter. In the future studies, researchers might consider individual characters of music sensitivity, training history or even music background (cultural style) and music preference as factors. This thesis suggested music could successfully be used to help memory encoding – or hinder it - and identified how music traits and music training alike could push that influence around. These new insights set me and others up for future studies on how we might using music as a memory aid, especially in clinical settings.

APPENDIX A. SUPPLEMENTORY STATISTICAL RESULTS

Table 2 Repeated measures ANOVA results

visual accuracy ~ Regularity*LevelOfMusicFamiliarity*MusicTrainingExperience + Error (subject)

	Df	Sum of Squares	F Value	P Value
Level of Music Familiarity	1	0.43	5.059	0.0246 *
Regularity Condition	2	0.14	0.883	0.4348
MusicTraining	1	0.05	0.536	0.4643
Level x Regularity	2	0.65	3.784	0.0229 *
Level x MusicTraining	1	0.04	0.455	0.5003
Regularity x MusicTrning	2	0.05	0.319	0.7272
Lvl x Regularity x MusicTrn	2	0.53	3.119	0.0445
Residuals	1751	150.08		

reaction time~ Regularity*LevelOfMusicFamiliarity*MusicTrainingExperience + Error (subject)

	Df	Sum of Squares	F Value	P Value
Level of Music Familiarity	1	1	0.303	0.5818
Regularity Condition	2	45	5.661	0.0035 **
MusicTraining	1	28	7.13	0.0077 **
Level x Regularity	2	69	8.702	0.0002 ***
Level x MusicTraining	1	4	1.074	0.3002
Regularity x MusicTrning	2	3	0.415	0.6602
Lvl x Regularity x MusicTrn	2	3	0.386	0.678
Residuals	1751	6922		

successful encoding phase ~ Regularity*LevelOfMusicFamiliarity*MusicTrainingExperience + Error (subject))

	Df	Sum of Squares	F Value	P Value
Level of Music Familiarity	1	4.2	2.479	0.1155
Regularity Condition	2	21.1	6.163	0.0022 **
Music Training	1	1.8	1.081	0.2985
Level x Regularity	2	30	8.775	0.0002 ***
Level x MusicTraining	1	0.7	0.394	0.5305
Regularity x MusicTrning	2	5.1	1.495	0.2245
Lvl x Regularity x MusicTrn	2	3.4	0.986	0.3732
Residuals	1715	2930.2		

Three main ANOVA tests were conducted for the main dataset. The tables showed the statistical results of the effects of music regularity, level of music familiarity and subject music training level on 1) visual sequences retrieval accuracy during retrieval task 2) reaction time for correct retrieved trials during retrieval task 3) successful encoding phase of each sequence during visual encoding/learning with subject as repeated measures. **(p<0.001: ***, p<0.01: **, p <0.05: *, 0.05< p<0.1: ·)**

Table 3 Post-hoc Tukey HSD test on visual retrieval accuracy

			95% Confidence Interval		
	Comparison	Mean Difference	Lower	Upper	p adj
**	Unlearned_Irregular - Learned_Irregular	0.0896	0.0157	0.1635	0.0073
	Learned_Control - Learned_Irregular	0.0473	-0.0307	0.1253	0.5115
	Unlearned_Control - Learned_Irregular	0.0609	-0.0171	0.139	0.2252
.	Learned_Regular - Learned_Irregular	0.0809	-0.0025	0.1642	0.0631
*	Unlearned_Regular - Learned_Irregular	0.0762	0.0015	0.151	0.0426
	Learned_Control - Unlearned_Irregular	-0.0422	-0.1064	0.0219	0.4162
	Unlearned_Control-Unlearned_Irregular	-0.0286	-0.0928	0.0355	0.7998
	Learned_Regular - Unlearned_Irregular	-0.0087	-0.0792	0.0618	0.9993
	Unlearned_Regular - Unlearned_Irregular	-0.0133	-0.0735	0.0468	0.9886
	Unlearned_Control - Learned_Control	0.0136	-0.0553	0.0825	0.9933
	Learned_Regular - Learned_Control	0.0335	-0.0413	0.1084	0.7971
	Unlearned_Regular - Learned_Control	0.0289	-0.0363	0.0941	0.804
	Learned_Regular - Unlearned_Control	0.0199	-0.0549	0.0948	0.9741
	Unlearned_Regular - Unlearned_Control	0.0153	-0.0499	0.0805	0.9852
	Unlearned_Regular - Learned_Regular	-0.0046	-0.0761	0.0668	1

The interactive effect between music regularity and level of familiarity on visual sequence retrieval accuracy was found to be significant. Thus, post-hoc pairwise comparison was conducted on all possible pairs of conditions. The table showed all statistical results from the Tukey HSD test comparing the mean retrieval accuracy between each pair of conditions. P value was adjusted and all significant pairs were highlighted. **($p<0.001$: ***, $p<0.01$: **, $p<0.05$: *, $0.05< p<0.1$: ·)**

Table 4 Post-hoc Tukey HSD test on retrieval RT

	Comparison	Mean Difference	95% Confidence Interval Lower	Upper	p adj
	Unlearned_Irregular - Learned_Irregular	-0.3844	-0.8866	0.1179	0.246
	Learned_Control - Learned_Irregular	0.2298	-0.3006	0.7602	0.8192
	Unlearned_Control - Learned_Irregular	0.0231	-0.5073	0.5535	1
	Learned_Regular - Learned_Irregular	-0.356	-0.9226	0.2105	0.4706
	Unlearned_Regular - Learned_Irregular	0.2038	-0.3045	0.7121	0.8629
***	Learned_Control - Unlearned_Irregular	0.6141	0.1778	1.0505	0.0009
.	Unlearned_Control-Unlearned_Irregular	0.4075	-0.0289	0.8438	0.0631
	Learned_Regular - Unlearned_Irregular	0.0283	-0.4513	0.5079	1
***	Unlearned_Regular - Unlearned_Irregular	0.5882	0.179	0.9973	0.0006
	Unlearned_Control - Learned_Control	-0.2067	-0.6751	0.2618	0.8075
*	Learned_Regular - Learned_Control	-0.5858	-1.0948	-0.0768	0.0134
	Unlearned_Regular - Learned_Control	-0.026	-0.4692	0.4172	1
	Learned_Regular - Unlearned_Control	-0.3791	-0.8881	0.1299	0.2748
	Unlearned_Regular - Unlearned_Control	0.1807	-0.2625	0.6239	0.8544
*	Unlearned_Regular - Learned_Regular	0.5598	0.0739	1.0457	0.0132

The interactive effect between music regularity and level of familiarity on reaction time of correctly retrieved sequences was found to be significant. Thus, post-hoc pairwise comparison was conducted on all possible pairs of conditions. The table showed all statistical results from the Tukey HSD test comparing the mean retrieval accuracy between each pair of conditions. P value was adjusted and all significant pairs were highlighted. **(p<0.001: ***, p<0.01: **, p <0.05: *, 0.05< p<0.1: ·)**

Table 5 Post-hoc Tukey HSD test on successful encoding phase

	Comparison	Mean Difference	95% Confidence Interval Lower	Upper	p adj
***	Unlearned_Irregular - Learned_Irregular	-0.496	-0.8277	-0.1642	0.0003
***	Learned_Control - Learned_Irregular	-0.5259	-0.8761	-0.1758	0.0003
***	Unlearned_Control - Learned_Irregular	-0.637	-0.9872	-0.2869	3.49E-06
***	Learned_Regular - Learned_Irregular	-0.6155	-0.9883	-0.2427	3.93E-05
***	Unlearned_Regular - Learned_Irregular	-0.431	-0.7669	-0.0951	0.0035
	Learned_Control - Unlearned_Irregular	-0.03	-0.3198	0.2599	0.9997
	Unlearned_Control-Unlearned_Irregular	-0.1411	-0.4309	0.1488	0.734
	Learned_Regular - Unlearned_Irregular	-0.1196	-0.4364	0.1973	0.8908
	Unlearned_Regular - Unlearned_Irregular	0.065	-0.2075	0.3374	0.9842
	Unlearned_Control - Learned_Control	-0.1111	-0.4218	0.1996	0.9113
	Learned_Regular - Learned_Control	-0.0896	-0.4256	0.2464	0.9739
	Unlearned_Regular - Learned_Control	0.0949	-0.1997	0.3895	0.9417
	Learned_Regular - Unlearned_Control	0.0215	-0.3145	0.3576	1
	Unlearned_Regular - Unlearned_Control	0.206	-0.0886	0.5006	0.3453
	Unlearned_Regular - Learned_Regular	0.1845	-0.1367	0.5057	0.5727

The interactive effect between music regularity and level of familiarity on successful encoding phases of visual sequences during learning was found to be significant. Thus, post-hoc pairwise comparison was conducted on all possible pairs of conditions. The table showed all statistical results from the Tukey HSD test comparing the mean retrieval accuracy between each pair of conditions. P value was adjusted and all significant pairs were highlighted. **(p<0.001: ***, p<0.01: **, p <0.05: *, 0.05< p<0.1: ·)**

Table 6 Pairwise comparison between conditions on amount of acquired sequences during each run of encoding phase (all subjects)

Run# of encoding	***p value of T-test between***	Control	Learned_Irregular	Learned_Regular	Unlearned_Irregular
1	Learned_Irregular	7.91E-05			
	Learned_Regular	0.837189543706179	0.00110958092313808		
	Unlearned_Irregular	0.416030884800693	0.00553725702347078	0.601270891499406	
	Unlearned_Regular	0.279605678639293	0.0107813180137895	0.451074093420472	0.816106348724823
2	Learned_Irregular	3.55E-05			
	Learned_Regular	0.770419884287599	0.000125361407475211		
	Unlearned_Irregular	0.322692787495609	0.00519541788834747	0.269351180750721	
	Unlearned_Regular	0.0701727992101076	0.0355149733306078	0.0699598150497192	0.474380537424139
3	Learned_Irregular	0.000537899609552845			
	Learned_Regular	0.382934316687753	0.000192718415889506		
	Unlearned_Irregular	0.933728573038017	0.00315018636827343	0.407333452370234	
	Unlearned_Regular	0.222538523937955	0.04553530305711	0.0711976802687342	0.324656230688856
4	Learned_Irregular	0.00182044227760516			
	Learned_Regular	0.682326906667546	0.00230092948395312		
	Unlearned_Irregular	0.554011130554345	0.00131369831286126	0.877032308366657	
	Unlearned_Regular	0.86995283433154	0.00976138303044962	0.619854804953315	0.512999066737513
5	Learned_Irregular	0.0225010423495961			
	Learned_Regular	0.826141908404693	0.0304527360439485		
	Unlearned_Irregular	0.594904863578285	0.0148466780105105	0.789047998173163	
	Unlearned_Regular	0.561583706101212	0.133914570164399	0.489920163003923	0.335709311640958

Red: p<0.05
Orange: trending, 0.05 <p < 0.1

Use this table as reference to Figure 6, here I compared all possible pairs of conditions across subjects on their mean of acquired visual sequences during each stage/run of encoding using t-test. The table showed all the p values of each t test between pairs. Due to small sample size after dividing data into conditions and because of the intention to investigate the trending patterns of effects, p-value was not adjusted.

Table 7 Pairwise comparison within Cluster 1 between conditions on amount of acquired sequences during each run of encoding phase

Run# of encoding	***p value of T-test between***	Control	Learned_Irregular	Learned_Regular	Unlearned_Irregular
1	Learned_Irregular	0.709224052			
	Learned_Regular	0.406858051	0.691870865		
	Unlearned_Irregular	0.762612346	0.951036322	0.647210016	
	Unlearned_Regular	0.078121867	0.065049339	0.025779617	0.074265717
2	Learned_Irregular	0.717219564			
	Learned_Regular	0.410019965	0.304993806		
	Unlearned_Irregular	0.794224628	0.930010489	0.347952899	
	Unlearned_Regular	0.188291462	0.407068168	0.065256044	0.359452269
3	Learned_Irregular	0.462738937			
	Learned_Regular	0.924412779	0.579431741		
	Unlearned_Irregular	0.81133512	0.399736137	0.772687673	
	Unlearned_Regular	0.562993077	0.892238415	0.675236337	0.479436832
4	Learned_Irregular	0.498237673			
	Learned_Regular	0.413390073	0.902641015		
	Unlearned_Irregular	0.545286597	0.26780429	0.218979087	
	Unlearned_Regular	0.521049528	0.975370223	0.878244262	0.281244576
5	Learned_Irregular	0.320296805			
	Learned_Regular	0.731271905	0.572192051		
	Unlearned_Irregular	0.572796303	0.178438643	0.432400311	
	Unlearned_Regular	0.510702725	0.76991299	0.785109655	0.290917857

Red: $p<0.05$
Orange: trending, $0.05 < p < 0.1$

In order to compare cluster 1 and cluster 3 learning behaviors and to detect the timing of the differences between conditions occurred, here I compared all possible pairs of conditions across subjects within Cluster 1 on their mean of acquired visual sequences during each stage/run of encoding using t-test. The table showed all the p values of each t test between pairs. Due to small sample size after dividing data into conditions plus clusters and because here my intention was to investigate the trending of effects, p-value was not adjusted.

Table 8 Pairwise comparison within Cluster 3 between conditions on amount of acquired sequences during each run of encoding phase

Run# of encoding	***p value of T-test between***	Control	Learned_Irregular	Learned_Regular	Unlearned_Irregular
1	Learned_Irregular	4.78E-10			
	Learned_Regular	0.328705332	1.81E-06		
	Unlearned_Irregular	0.134380148	1.22E-05	0.649259176	
	Unlearned_Regular	0.69993072	1.85E-07	0.607683705	0.333657243
2	Learned_Irregular	9.09E-06			
	Learned_Regular	0.458529647	8.27E-06		
	Unlearned_Irregular	0.399850888	0.001346409	0.171456104	
	Unlearned_Regular	0.13630884	0.007593814	0.054368359	0.571724424
3	Learned_Irregular	0.004828975			
	Learned_Regular	0.094927468	0.000182304		
	Unlearned_Irregular	0.641777403	0.039391585	0.077745607	
	Unlearned_Regular	0.488056352	0.061939436	0.050261202	0.843030121
4	Learned_Irregular	0.01687318			
	Learned_Regular	0.192538524	0.001577604		
	Unlearned_Irregular	0.764191837	0.019937462	0.383593027	
	Unlearned_Regular	0.655385964	0.014288539	0.456278168	0.899085515
5	Learned_Irregular	0.493143518			
	Learned_Regular	0.378553936	0.17613531		
	Unlearned_Irregular	0.542983216	0.263400904	0.812985609	
	Unlearned_Regular	0.847365344	0.447328332	0.550726935	0.718539685

Red: $p<0.05$
Orange: trending, $0.05 < p < 0.1$

In order to compare cluster 1 and cluster 3 learning behaviors and to detect the timing of the differences between conditions occurred, here I compared all possible pairs of conditions across subjects within Cluster 3 on their mean of acquired visual sequences during each stage/run of encoding using t-test. The table showed all the p values of each t test between pairs. Due to small sample size after dividing data into conditions plus clusters and because here my intention was to investigate the trending of effects, p-value was not adjusted.

Table 9 Pairwise comparison between conditions on retrieval task results (retrieval accuracy + RT for correct trials) within each cluster

Cluster # 1: retrieval accuracy

	Learned_Irregular	Learned_Control	Learned_Regular	Unlearned_Irregular	Unlearned_Control
Learned_Control	0.447293234263794				
Learned_Regular	0.119335176583814	0.372100200931237			
Unlearned_Irregular	0.0332035811296477	0.143100186563557	0.648384843394718		
Unlearned_Control	0.619797915576247	0.774348974675903	0.243447160405029	0.0773361574039694	
Unlearned_Regular	0.122033961221831	0.403590089230433	0.90034604186466	0.523690902325737	0.257713621941817

Cluster # 1: reaction time

	Learned_Irregular	Learned_Control	Learned_Regular	Unlearned_Irregular	Unlearned_Control
Learned_Control	0.894755479369697				
Learned_Regular	0.521201130278596	0.401485939085672			
Unlearned_Irregular	0.265828056047203	0.164985487150061	0.658029494157124		
Unlearned_Control	0.622945561495848	0.693497287260159	0.223650941926687	0.0712632614697523	
Unlearned_Regular	0.423781299384516	0.464147289161076	0.115137421118164	0.0250651409802417	0.749605893555242

Cluster # 3: retrieval accuracy

	Learned_Irregular	Learned_Control	Learned_Regular	Unlearned_Irregular	Unlearned_Control
Learned_Control	0.408130541901174				
Learned_Regular	0.0392772585482652	0.156204012303223			
Unlearned_Irregular	0.0449417638054481	0.199653578574153	0.737638672965689		
Unlearned_Control	0.0407473049564503	0.173727818452295	0.864135105565267	0.864867238339617	
Unlearned_Regular	0.0575447726250283	0.243251540980403	0.668689109775562	0.909993257622691	0.784308636918746

Cluster # 3: reaction time

	Learned_Irregular	Learned_Control	Learned_Regular	Unlearned_Irregular	Unlearned_Control
Learned_Control	0.320705185549117				
Learned_Regular	0.849904545396648	0.208985537950107			
Unlearned_Irregular	0.267159303482622	0.0131895973978339	0.340878588338222		
Unlearned_Control	0.742806871500681	0.444240941229123	0.576013962216883	0.0899820773913985	
Unlearned_Regular	0.146642842363564	0.64181561783788	0.0782673347854371	0.00144412436767119	0.19631416991661

Red: $p<0.05$
Orange: trending, $0.05 < p < 0.1$

In order to apply the clustering results from encoding to retrieval phase and compare cluster1 and cluster3's retrieval performance, here I separately ran t-test and compared all possible pairs of conditions across subjects within each cluster on their mean retrieval accuracy and RT for correct trials. The table showed all the p values of each t test between pairs unadjusted. Significant and trending differences were highlighted.

Table 10 Repeated measures ANOVA results for follow-up baseline task

visual accuracy ~ Regularity*LevelOfMusicMemory + Error (subject)

	Df	Sum of Squares	F Value	P Value
Level of Music Familiarity	1	0	0.003	0.959
Regularity Condition	2	0.04	0.303	0.739
Level x Regularity	2	0.05	0.426	0.653
Residuals	695	42.3		

reaction time~ Regularity*LevelOfMusicMemory + Error (subject)

	Df	Sum of Squares	F Value	P Value
Level of Music Familiarity	1	0.1	0.017	0.895
Regularity Condition	2	4	0.602	0.5481
Level x Regularity	2	19.5	2.921	0.0546 *
Residuals	695	2190.2		

successful encoding phase ~ Regularity*LevelOfMusicMemory + Error (subject)

	Df	Sum of Squares	F Value	P Value
Level of Music Familiarity	1	2.3	1.433	0.2317
Regularity Condition	2	10.1	3.209	0.041 *
Level x Regularity	2	7.4	2.359	0.0953 .
Residuals	713	1120.9		

Same ANOVA tests to the main task were conducted for the follow-up dataset. Due to the small sample size, dividing subjects into groups with different music training will lead to less test power, I only tested the effects of music regularity and level of music familiarity and their interactive effects on 1) visual sequences retrieval accuracy during retrieval task 2) reaction time for correct retrieved trials during retrieval task 3) successful encoding phase of each sequence during visual encoding/learning with subject as repeated measures. **($p<0.001$: ***, $p<0.01$: **, $p <0.05$: *, $0.05< p<0.1$: ·)**

REFERENCES

Albert, M. L., Sparks, R. W., & Helm, N. A. (1973). Melodic intonation therapy for aphasia. *Archives of Neurology*, *29*(2), 130–131.

Alexander, W. H., & Brown, J. W. (2014). A general role for medial prefrontal cortex in event prediction. *Frontiers in Computational Neuroscience*, *8*, 69.

Barnett, A. J., O'Neil, E. B., Watson, H. C., & Lee, A. C. H. (2014). The human hippocampus is sensitive to the durations of events and intervals within a sequence. *Neuropsychologia*, *64*, 1–12.

Belfi, A. M., Karlan, B., & Tranel, D. (2016). Music evokes vivid autobiographical memories. *Memory*, *24*(7), 979–989.

Bernatzky, G., Bernatzky, P., Hesse, H.-P., Staffen, W., & Ladurner, G. (2004). Stimulating music increases motor coordination in patients afflicted with Morbus Parkinson. *Neuroscience Letters*, *361*(1–3), 4–8.

Bigand, E., Delbé, C., Poulin-Charronnat, B., Leman, M., & Tillmann, B. (2014). Empirical evidence for musical syntax processing? Computer simulations reveal the contribution of auditory short-term memory. *Frontiers in Systems Neuroscience*, *8*, 94.

Boltz, M., Schulkind, M., & Kantra, S. (1991). Effects of background music on the remembering of filmed events. *Memory & Cognition*, *19*(6), 593–606.

Butler, D. (1989). Describing the perception of tonality in music: A critique of the tonal hierarchy theory and a proposal for a theory of intervallic rivalry. *Music Perception*, *6*(3), 219–241.

Chou, P. T.-M. (2010). Attention Drainage Effect: How Background Music Effects Concentration in Taiwanese College Students. *Journal of the Scholarship of Teaching and Learning*, *10*(1), 36–46.

Cross, I. (2001). Music, cognition, culture, and evolution. *Annals of the New York Academy of Sciences*, *930*(1), 28–42.

Cuddy, L. L., Sikka, R., Silveira, K., Bai, S., & Vanstone, A. (2017). Music-evoked autobiographical memories (MEAMs) in Alzheimer disease: Evidence for a positivity effect. *Cogent Psychology*, *4*(1), 1277578.

Curran, T., & Keele, S. W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(1), 189.

Degé, F., Wehrum, S., Stark, R., & Schwarzer, G. (2011). The influence of two years of school music training in secondary school on visual and auditory memory. *European Journal of Developmental Psychology*, *8*(5), 608–623.

Dibben, N. (1994). The cognitive reality of hierarchic structure in tonal and atonal music. *Music Perception*, *12*(1), 1–25.

Eichenbaum, H. (2013). Memory on time. *Trends in Cognitive Sciences*, *17*(2), 81–88.

Eichenbaum, H. (2014). Time cells in the hippocampus: a new dimension for mapping memories. *Nature Reviews Neuroscience*, *15*(11), 732–744.

Eichenbaum, H., & Fortin, N. (2003). Episodic memory and the hippocampus: It's about time. *Current Directions in Psychological Science*, *12*(2), 53–57.

Ettlinger, M., Margulis, E. H., & Wong, P. (2011). Implicit memory in music and language. *Frontiers in Psychology*, *2*, 211.

Falkenstein, M. (1990). Effects of errors in choice reaction tasks on the ERP under focused and divided attention. *Psychophysiological Brain Research*.

Falkenstein, Michael, Hielscher, H., Dziobek, I., Schwarzenau, P., Hoormann, J., Sundermann, B., & Hohnsbein, J. (2001). Action monitoring, error detection, and the basal ganglia: an ERP study. *Neuroreport*, *12*(1), 157–161.

Ferreri, L., Aucouturier, J., Muthalib, M., Bigand, E., & Bugaiska, A. (2013). Music improves verbal memory encoding while decreasing prefrontal cortex activity: an fNIRS study. *Frontiers in Human Neuroscience*, *7*, 779.

Furnham, A., & Bradley, A. (1997). Music while you work: The differential distraction of background music on the cognitive test performance of introverts and extraverts. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *11*(5), 445–455.

Furnham, A., & Strbac, L. (2002). Music is as distracting as noise: The differential distraction of background music and noise on the cognitive test performance of introverts and extraverts. *Ergonomics*, *45*(3), 203–217.

Gais, S., & Born, J. (2004). Declarative memory consolidation: mechanisms acting during human sleep. *Learning & Memory*, *11*(6), 679–685.

Gilbert, P. E., & Kesner, R. P. (2002). Role of rodent hippocampus in paired-associate learning involving associations between a stimulus and a spatial location. *Behavioral Neuroscience*, *116*(1), 63.

Gimbel, S. I., & Brewer, J. B. (2011). Reaction time, memory strength, and fMRI activity during memory retrieval: Hippocampus and default network are differentially responsive during recollection and familiarity judgments. *Cognitive Neuroscience*, *2*(1), 19–26.

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience*, *15*(6), 883–899.

Hasselmo, M. E. (2006). The role of acetylcholine in learning and memory. *Current Opinion in Neurobiology*, *16*(6), 710–715.

Havas, V., Taylor, J. S. H., Vaquero, L., de Diego-Balaguer, R., Rodríguez-Fornells, A., & Davis, M. H. (2018). Semantic and phonological schema influence spoken word learning and overnight consolidation. *Quarterly Journal of Experimental Psychology*, *71*(6), 1469–1481.

Henke, K., Weber, B., Kneifel, S., Wieser, H. G., & Buck, A. (1999). Human hippocampus associates information in memory. *Proceedings of the National Academy of Sciences*, *96*(10), 5884–5889.

Herholz, S. C., Lappe, C., & Pantev, C. (2009). Looking for a pattern: an MEG study on the abstract mismatch negativity in musicians and nonmusicians. *BMC Neuroscience*, *10*(1), 42.

Holdstock, J. S., Crane, J., Bachorowski, J.-A., & Milner, B. (2010). Equivalent activation of the hippocampus by face-face and face-laugh paired associate learning and recognition. *Neuropsychologia*, *48*(13), 3757–3771.

Huron, D. (2001). Is music an evolutionary adaptation? *Annals of the New York Academy of Sciences*, *930*(1), 43–61.

Irish, M., Cunningham, C. J., Walsh, J. B., Coakley, D., Lawlor, B. A., Robertson, I. H., & Coen, R. F. (2006). Investigating the enhancing effect of music on autobiographical memory in mild Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, *22*(1), 108–120.

Janata, P. (2009). The neural architecture of music-evoked autobiographical memories. *Cerebral Cortex*, *19*(11), 2579–2594.

Janata, P., & Grafton, S. T. (2003). Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience*, *6*(7), 682–687.

Jäncke, L., & Sandmann, P. (2010). Music listening while you learn: No influence of background music on verbal learning. *Behavioral and Brain Functions*, *6*(1), 1–14.

Jonides, J. (2008). Musical skill and cognition. *Learning, Arts, and the Brain*, 11.

Kang, H. J., & Williamson, V. J. (2014). Background music can aid second language learning. *Psychology of Music*, *42*(5), 728–747.

Kasdan, A., & Kiran, S. (2018). Please don't stop the music: Song completion in patients with aphasia. *Journal of Communication Disorders*, *75*, 72–86.

Kesner, R. P., Gilbert, P. E., & Barua, L. A. (2002). The role of the hippocampus in memory for the temporal order of a sequence of odors. *Behavioral Neuroscience*, *116*(2), 286.

Knutson, B., Fong, G. W., Bennett, S. M., Adams, C. M., & Hommer, D. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. *Neuroimage*, *18*(2), 263–272.

Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of Cluster in K-Means Clustering. *International Journal*, *1*(6), 90–95.

Koelsch, S., Jentschke, S., Sammler, D., & Mietchen, D. (2007). Untangling syntactic and sensory processing: An ERP study of music perception. *Psychophysiology*, *44*(3), 476–490.

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences*, *110*(38), 15443–15448.

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences*, *23*(1), 63–77.

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin*, *126*(1), 159.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, *5*(4), 579.

Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural networks for beat perception in musical rhythm. *Frontiers in Systems Neuroscience*, *9*, 159.

Lartillot, O., Eerola, T., Toiviainen, P., & Fornari, J. (2008). Multi-Feature Modeling of Pulse Clarity: Design, Validation and Optimization. *ISMIR*, 521–526.

Lartillot, O., & Toiviainen, P. (2007). A Matlab toolbox for musical feature extraction from audio. *International Conference on Digital Audio Effects*, *237*, 244.

Leman, M. (2012). *Music and schema theory: Cognitive foundations of systematic* *musicology* (Vol. 31). Springer Science & Business Media.

Lerdahl, F., & Jackendoff, R. (1983). An overview of hierarchical structure in music. *Music Perception*, *1*(2), 229–252.

Levitin, D. J. (2006). *This is your brain on music: The science of a human obsession*. Penguin.

Lord, T. R., & Garner, J. E. (1993). Effects of music on Alzheimer patients. *Perceptual and Motor Skills*, *76*(2), 451–455.

McKenzie, S., Frank, A. J., Kinsky, N. R., Porter, B., Rivière, P. D., & Eichenbaum, H. (2014). Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. *Neuron*, *83*(1), 202–215.

McKenzie, S., Robinson, N. T. M., Herrera, L., Churchill, J. C., & Eichenbaum, H. (2013). Learning causes reorganization of neuronal firing patterns to represent related experiences within a hippocampal schema. *Journal of Neuroscience*, *33*(25), 10243–10256.

Mencke, I., Omigie, D., Wald-Fuhrmann, M., & Brattico, E. (2019). Atonal music: can uncertainty lead to pleasure? *Frontiers in Neuroscience*, *12*, 979.

Menon, V., & Levitin, D. J. (2005). The rewards of music listening: response and physiological connectivity of the mesolimbic system. *Neuroimage*, *28*(1), 175–184.

Moore, K. S. (2013). A systematic review on the neural effects of music on emotion regulation: implications for music therapy practice. *Journal of Music Therapy*, *50*(3), 198–242.

Moussard, A., Bigand, E., Belleville, S., & Peretz, I. (2012). Music as an aid to learn new verbal information in Alzheimer's disease. *Music Perception: An Interdisciplinary Journal*, *29*(5), 521–531.

Nadel, L., Samsonovich, A., Ryan, L., & Moscovitch, M. (2000). Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results. *Hippocampus*, *10*(4), 352–368.

Nakamura, S., Sadato, N., Oohashi, T., Nishina, E., Fuwamoto, Y., & Yonekura, Y. (1999). Analysis of music–brain interaction with simultaneous measurement of regional cerebral blood flow and electroencephalogram beta rhythm in human subjects. *Neuroscience Letters*, *275*(3), 222–226.

Negus, K. (2013). *Music genres and corporate cultures*. Routledge.

Norton, A., Winner, E., Cronin, K., Overy, K., Lee, D. J., & Schlaug, G. (2005). Are there pre-existing neural, cognitive, or motoric markers for musical ability? *Brain and Cognition*, *59*(2), 124–134.

Orio, N., & Schwarz, D. (2001). *Alignment of monophonic and polyphonic music to a score*.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6*(7), 674–681.

Patel, A. D. (2010). *Music, language, and the brain*. Oxford university press.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, *10*(6), 717–733.

Preston, A. R., & Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Current Biology*, *23*(17), R764–R773.

Proverbio, A. M., Nasi, V. L., Arcari, L. A., De Benedetto, F., Guardamagna, M., Gazzola, M., & Zani, A. (2015). The effect of background music on episodic memory and autonomic responses: listening to emotionally touching music enhances facial memory capacity. *Scientific Reports*, *5*, 15219.

Pulvermüller, F., & Assadollahi, R. (2007). Grammar or serial order?: Discrete combinatorial brain mechanisms reflected by the syntactic mismatch negativity. *Journal of Cognitive Neuroscience*, *19*(6), 971–980.

Raab, D. H. (1962). Division of psychology: Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, *24*(5 Series II), 574–590.

Rainey, D. W., & Larsen, J. D. (2002). The effect of familiar melodies on initial learning and long-term memory for unconnected text. *Music Perception*, *20*(2), 173–186.

Reaves, S., Graham, B., Grahn, J., Rabannifard, P., & Duarte, A. (2016). Turn off the music! Music impairs visual associative memory performance in older adults. *The Gerontologist*, *56*(3), 569–577.

Robinson, M. D., Johnson, J. T., & Herndon, F. (1997). Reaction time and assessments of cognitive effort as predictors of eyewitness memory accuracy and confidence. *Journal of Applied Psychology*, *82*(3), 416.

Roden, I., Könen, T., Bongard, S., Frankenberg, E., Friedrich, E. K., & Kreutz, G. (2014). Effects of music training on attention, processing speed and cognitive music abilities—findings from a longitudinal study. *Applied Cognitive Psychology*, *28*(4), 545–557.

Rohrmeier, M. A., & Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology*, *83*(2), 164–175.

Rohrmeier, M., & Rebuschat, P. (2012). Implicit learning and acquisition of music. *Topics in Cognitive Science*, *4*(4), 525–553.

Salimpoor, V. N., van den Bosch, I., Kovacevic, N., McIntosh, A. R., Dagher, A., & Zatorre, R. J. (2013). Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science*, *340*(6129), 216–219.

Samson, S., & Zatorre, R. J. (1992). Learning and retention of melodic and verbal information after unilateral temporal lobectomy. *Neuropsychologia*, *30*(9), 815–826.

Schlaug, G., Norton, A., Marchina, S., Zipse, L., & Wan, C. Y. (2010). From singing to speaking: facilitating recovery from nonfluent aphasia. *Future Neurology*, *5*(5), 657–665.

Seashore, C. E. (1937). The psychology of music. *Music Educators Journal*, *23*(4), 30–33.

Smith, S. M. (1985). Background music and context-dependent memory. *The American Journal of Psychology*, 591–603.

Tan, S.-L., Pfordresher, P., & Harré, R. (2017). *Psychology of music: From sound to significance*. Routledge.

Thompson, W. F., Schellenberg, E. G., & Letnic, A. K. (2012). Fast and loud background music disrupts reading comprehension. *Psychology of Music*, *40*(6), 700–708.

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., Witter, M. P., & Morris, R. G. M. (2007). Schemas and memory consolidation. *Science*, *316*(5821), 76–82.

Van Kesteren, M. T. R., Brown, T. I., & Wagner, A. D. (2018). Learned spatial schemas and prospective hippocampal activity support navigation after one-shot learning. *Frontiers in Human Neuroscience*, *12*, 486.

Van Kesteren, M. T. R., Fernández, G., Norris, D. G., & Hermans, E. J. (2010). Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proceedings of the National Academy of Sciences*, *107*(16), 7550–7555.

van Kesteren, M. T. R., Rijpkema, M., Ruiter, D. J., Morris, R. G. M., & Fernández, G. (2014). Building on prior knowledge: schema-dependent encoding processes relate to academic performance. *Journal of Cognitive Neuroscience*, *26*(10), 2250–2261.

Van Kesteren, M. T. R., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, *35*(4), 211–219.

Wallace, W. T. (1994). Memory for music: Effect of melody on recall of text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(6), 1471.

Watanabe, T., Yagishita, S., & Kikyo, H. (2008). Memory of music: roles of right hippocampus and left inferior frontal gyrus. *Neuroimage*, *39*(1), 483–491.

Wedin, L. (1972). A multidimensional study of perceptual-emotional qualities in music. *Scandinavian Journal of Psychology*, *13*(1), 241–257.

Weiermann, B., Cock, J., & Meier, B. (2010). What matters in implicit task sequence learning: perceptual stimulus features, task sets, or correlated streams of information? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(6), 1492.

Yalch, R. F. (1991). Memory in a jingle jungle: Music as a mnemonic device in communicating advertising slogans. *Journal of Applied Psychology*, *76*(2), 268.