Expertise effects on visual change detection in the music reading domain: evidence from eye movements

Abigail L. Kleinsmith
University at Albany, State University of New York, kleinsmithabigail@gmail.com

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etz

Part of the Music Commons, and the Psychology Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive.
Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
Expertise Effects on Visual Change Detection in the Music Reading Domain:

Evidence from Eye Movements

by

Abigail L. Kleinsmith

A Dissertation

Submitted to the University at Albany, State University of New York

In Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

College of Arts and Sciences

Department of Psychology

2022
Abstract

Theoretical perspectives in the chess expertise literature, such as chunking and template theories, assume that experts acquire the ability to process domain-specific visual features as larger patterns. Eye tracking techniques can test predictions derived from these theories, because the eye movement record provides fine-grained information about where and when experts are looking during a domain-specific task. In this dissertation, I assessed the generalizability of chunking and template theories to the domain of music reading expertise with a novel music-related variant of the flicker paradigm. Across two experiments, I monitored the eye movements of 60 expert musicians (with at least 10 years of music reading experience) and 60 non-musicians (who could not read music) while they located a single music note or note stack that was rapidly appearing and disappearing within a larger music score. In Experiment 1, I manipulated visual complexity to test the prediction that experts’ domain-specific knowledge permits them to compensate for the processing difficulty produced by increases in the visual complexity of the music scores. In support of this prediction, experts completed trials with more efficient eye movements than non-musicians. Also, the complex trials elicited shorter saccade amplitudes and longer fixation durations than the simple trials, and this complexity effect was larger for non-musicians than experts. In Experiment 2, I manipulated familiarity by contrasting upright (i.e., typically oriented) scores with scores rotated by $90^\circ$ to test the prediction that visual expertise is perceptually specific. In support of this prediction, I observed no expertise effects on eye movements in the rotated condition, while observing robust expertise effects in the upright condition. Specifically, expertise effects were eliminated in the rotated condition, and the experts made numerically more fixations than non-musicians during rotated trials. Taken together, these findings support the conclusions that music reading experts: (a) can rely on domain-specific knowledge to compensate for increases in task difficulty, and (b) display perceptual advantages that are remarkably specific to familiar contexts. Consistent with the assumptions of chunking and template theories, and similar findings in
other domains of expertise like chess and medicine, these results suggest that expert musicians can use chunking mechanisms (e.g., perceptual grouping) to facilitate encoding during domain-specific change detection tasks. This work helps to clarify the boundary conditions under which perceptual expertise effects occur and assesses the generalizability of the chunking and template framework within a novel domain of expertise.

*Keywords:* eye tracking; music reading expertise; flicker paradigm; change detection; chunking and template theories
Dedication

For my students: past, present, and future.
Acknowledgements

I would not exist today without the support of others.
This work would not exist without their presence in my life.
Thank you...

to my committee,
Heather, Dominique, and Jeanette: for your invaluable support, for your thoughtful and compassionate feedback on my work, and for so generously sharing your expertise with me.

to my family,
Gretchen: for being my first best friend.
Kym and Kelly: for always being my cheerleaders.
Vincenzo: for your incredible partnership and commitment.
Karen and Nick: for opening your home and your hearts to me.
Cindy and John: for being “Mom and Dad” when I needed it the most.

to my friends,
Jenny: for being my truest steadfast companion. A decade later, we are still standing.
Jeremiah: for being the brother I always wanted.
Rita and Jason: for accepting me and loving me as I am.
Allison, Stephanie, and Deana: for sharing your journey with me.
Rachel and Chris: for teaching me the joys of true friendship and bibimbap.

to my teachers,
David: for opening my brain, mind, and consciousness, and for being my first true mentor.
John: for teaching me that my brain is wider than the sky.
Juan: for teaching me to play the music after I learned the notes.
Jim: for the love of cognitive psychology.
Tram: for teaching me the art of healthy skepticism as a scientist and critical thinker.
Bruce: for the love of statistics, and for Shiny apps.
Greg: for teaching me that I have always had the key to unlock my true potential.
Heather: for teaching me how to write, and most importantly, for your unrelenting patience.
to my colleagues,

  Casey, Kaitlyn, and Betty: for laughs, hugs, and camaraderie.
  Linda, Leah, Darlene, and Jenn: for always being a bright spot in my day.
  Peter: for coffee. Always coffee.

to my labmates, past and present,

  Jeff: for being my forever-five-percent.
  George: for long discussions and chilly walks on the podium.
  Kinrera and Eliza: for truly seeing me, and then supporting me even harder.
  Maddie and Nicole: for being the future of the Visual Cognition Lab.

to my research assistants, past and present,

  Becky: for your impeccable work with stimulus creation.
  Raina: for sharing your passion with me, and for beautifully coding our stimulus characteristics.
  Brian: for your expert data, and for OSPAN coding.
  Yadi, Isabelle, Kenya, jaSonia, Caity, Alec, Marisa, Nicole, Kaelyn, Allyson, Mehwish, Natalie, Daniela, Sabrina, Roy, Joslin, Camila, and Andrea: for years of exploration and hundreds of participants.

to my students, past and present,

  Gaby, Mariah, Kassidy, Katarina, and Eddie: for sharing your spark with me.
  Amy, Alex, Sally, Amen, Luke, Bailey, Eliza, and Sam: for living fully in your truth and for trusting me to see you.

to those who have passed,

  Keith: for the love of music.
  Cyndy: for the love of reading.
  Vivian and Clarence: for the most wholesome parts of my childhood.
  Ayla: for true unconditional love.
  Bob: for teaching me how to harmonize.
  Nathan: for being sunshine in the darkness.
# Table of Contents

Introduction ......................................................................................................................... 1

Chunking and Template Theories......................................................................................... 3

Eye Tracking .......................................................................................................................... 6

The Domain of Chess ........................................................................................................... 9

Perceptual Expertise Effects in the Domain of Chess .......................................................... 11

The Domains of Chess and Music Reading ........................................................................ 13

Perceptual Expertise Effects in the Domain of Music Reading ............................................ 14

The Change Detection Literature ....................................................................................... 17

The Dissertation .................................................................................................................. 19

Experiment 1 ....................................................................................................................... 23

Experiment 2 ....................................................................................................................... 32

General Discussion ............................................................................................................. 42

Limitations and Future Research Directions ..................................................................... 50

Broader Connections ......................................................................................................... 61

References ............................................................................................................................ 68

Tables .................................................................................................................................. 86

Figures .................................................................................................................................. 89

Appendix ............................................................................................................................... 96
List of Tables

Table 1. Mean and standard error of the mean (in parentheses) for the variables analyzed in Experiment 1: accuracy, reaction time (seconds), fixation duration (milliseconds), number of fixations, and saccade amplitude (degree of visual angle).

Table 2. Mean and standard error of the mean (in parentheses) for the variables analyzed in Experiment 2: accuracy, reaction time (seconds), fixation duration (milliseconds), number of fixations, and saccade amplitude (degree of visual angle), and OSPAN scores.

Table 3. OSPAN analysis for Experiment 2.
List of Figures

**Figure 1.** An illustration of the foveal, parafoveal, and peripheral regions of the visual field during the processing of a music score.

**Figure 2.** Example of potential chunks and templates in the chess domain and the music-reading domain.

**Figure 3.** Sample stimuli and a schematic illustration of a trial progression in Experiment 1.

**Figure 4.** Means and standard errors for accuracy, reaction time, fixation duration, fixation count, and saccade amplitude, as a function of expertise and complexity in Experiment 1.

**Figure 5.** Sample stimuli and a schematic illustration of a trial progression in Experiment 2.

**Figure 6.** Sample OSPAN trial in Experiment 2.

**Figure 7.** Means and standard errors for accuracy, reaction time, fixation duration, fixation count, and saccade amplitude, as a function of expertise and complexity in Experiment 2.
List of Appendices

Appendix. Demographic questionnaire used in Experiments 1 and 2.
Introduction

Eye tracking methodology has been used extensively to document the remarkable perceptual skill of experts in different domains, including chess (e.g., Reingold & Sheridan, 2011), medicine (e.g., Sheridan & Reingold, 2017), and face perception (e.g., Richler & Gauthier, 2014). This prior work highlights the usefulness of eye tracking methodology for exploring differences across levels of expertise (for reviews, see Madell & Hébert, 2008; Rayner, 2009). Experts’ enhanced perceptual skills permit efficient eye movement behavior during a domain-specific task, as evidenced by fewer fixations (periods of time when the eye is relatively still, and visual input is encoded) and longer saccades (periods of time when the eye is in motion, and visual input is suppressed) (for reviews, see Reingold & Sheridan, 2011; Sheridan & Reingold, 2017). An example of this efficiency is that experts are faster than non-experts at extracting relevant information from an image during a domain specific task (for a review, see Reingold & Sheridan, 2011). Experts can also encode more information in a single glance than can non-experts, as indexed by their larger perceptual span (or, visual span) (e.g., Reingold et al., 2001).

One of the main theoretical frameworks for interpreting these expertise effects within the domain of chess is that of chunking and template theories (Chase & Simon, 1973a, 1973b; Gobet & Simon, 1996, 2000). These theories assume that experts acquire large mental vocabularies of domain-specific information that is grouped (or “chunked”) by meaning (e.g., Chase & Simon, 1973a), and that experts make use of these chunks in conjunction with flexible mental templates that support and facilitate experts’ performance on a domain-related task (e.g., Gobet & Simon, 1996). Chunking and template theories predict that experts should be able to rely on mental structures acquired over thousands of hours of deliberate practice (e.g., Ericsson et al., 1993) to compensate for increases in domain-specific task complexity.
To clarify the extent to which these predictions are generalizable, as opposed to being tied
directly to a specific domain, task demand, or modality, extensions to novel domains (e.g., music reading
expertise) should be considered. In my dissertation, I discuss how the music reading domain is ideal for
examining the boundary conditions of these expertise effects. The music reading domain is complex,
dynamic, and uniquely multimodal, in addition to providing a number of methodological advantages
that will be described within the thesis. Critically, similarities between the domain of music reading
expertise and the domain of chess (where robust support for chunking and template theories is
observed), provide compelling reasons to investigate the question of whether chunking and template
theories can also constitute a theoretical framework for interpreting expertise effects within a novel
domain. For example, music notes and chess pieces can both be perceptually (i.e., visually) grouped into
meaningful clusters that could form “chunks” of information. Music scores and chess boards also both
contain meaningful larger structures that could form “templates”. Ultimately, exploring the
generalizability of the predictions of chunking and template theories contributes to ongoing efforts
within the field to develop more comprehensive and unified theories of expertise (for a related
discussion, see e.g., Brams et al., 2019).

In my dissertation, I will test whether the assumptions of chunking and template theories could
extend to the novel domain of music reading. First, I will describe the theoretical background of
chunking and template theories, as well as discuss why eye tracking methodology is a useful method for
testing the predictions generated by these theories. Next, I will provide a review of relevant evidence
within the perceptual expertise literature. This will include a discussion of the similarities between chess
and music that help set the precedent for exploring the specific predictions in my dissertation research.

---

1 As an example in music, familiarity with a specific compositional style (i.e., a *sonata*) facilitates rapid and accurate
encoding. See Figure 2 for more about the similarities between the domains of chess and music, in terms of
perceptual “chunks” and “templates”.
After this, I will focus my discussion of perceptual expertise effects to the change detection literature, and specifically highlight how expertise effects on change detection are inconsistent across many domains of expertise (e.g., driving, sports, and video games). This gap in the change detection literature is directly addressed with my dissertation research.

Next, I will describe the primary goals of my dissertation in addition to introducing the novel paradigm that I adapted from the domain of chess for my needs in the domain of music reading. I will then describe the two experiments of the dissertation. I will summarize the main findings of both experiments and describe the most important implications of this work to the field of expertise research. I will also provide a discussion of the strengths and limitations of the present work as well as suggest some next steps for future research directions. Finally, I will conclude with a speculative section about potential broader connections between my dissertation work and the music education literature.

**Chunking and Template Theories**

Chunking and template theories posit that experts’ perceptual advantages are acquired over thousands of hours of experience and practice within a specific domain. This concentrated practice results in the development of large mental vocabularies of domain-specific stimuli (i.e., chunks) that support experts’ superior problem-solving abilities when holding a domain-specific problem in short-term memory (Chase & Simon 1973a; 1973b). Importantly, this concentrated practice also results in the development of larger long-term memory structures (i.e., templates) that support and supplement experts’ short-term memory (Gobet & Simon, 1996; 2000). Taken together, chunking and template theories suggest that these short- and long-term memory structures work in conjunction to allow

---

2 Data presented in Experiment 1 are previously published (Sheridan & Kleinsmith, 2021).
experts to process domain-specific stimuli in terms of broader patterns, as opposed to their individual features.

The domain of chess expertise has supported the assumption of chunking and template theories that a perceptual advantage is a fundamental component of chess expertise (Chase & Simon 1973a; Gobet & Simon, 1996). Chase and Simon (1973a) observed that expert chess players have a remarkable ability to recreate chess board configurations from memory after extremely brief visual presentations, in contrast to less-skilled players who exhibited worse performance. In contrast, when chess pieces were presented in random configurations, both groups performed approximately equally well on this memory task (i.e., expert accuracy dropped to a similar level as the non-expert). In this way, disrupting the legality of chess piece placement greatly impacted (although did not completely eliminate\(^3\)) the experts’ performance advantages. This finding led Chase and Simon to conclude that experts were likely storing a large quantity of legal chess piece configurations in memory. These results form the basis of Chase and Simon’s (1973a) chunking theory, which posits that the perceptual advantages of experts can be attributed to memory structures comprised of domain-specific information, which are called “chunks”.

While chunking theory helps explain how experts perceive larger combinations of individual features than non-experts, Gobet and Simon (1996) noted that chunking theory alone could not sufficiently account for experts’ extraordinary performance. Specifically, Gobet and Simon suggested that chunking theory fails to adequately explain the role of experts’ short- and long-term memory, and the interaction between these two cognitive systems, during domain-specific tasks (1996). To address these limitations, Gobet and Simon (1996, 2000) developed an extension to chunking theory – namely,

---

\(^3\) See Gobet & Oliver (2016) for a discussion of the dissociability of “chunks” from “templates”, as well as the random occurrence of meaningful “chunks” within a larger random pattern.
Template theory – which suggests that experts group together individual chunks in long-term memory, which promotes efficiency and flexibility in completing domain-specific tasks.

Template theory suggests that groups of chunks form larger memory structures called templates. Templates are long-term memory structures that are similar to “schemas”, and templates contain domain-specific information about familiar patterns and larger collections of chunks. Templates include flexible slots into which both new and old chunks can be “inserted”. Slots storing older information help supplement short-term memory, as older chunks in long-term memory can be rapidly integrated with newly-learned patterns or frameworks (e.g., optimal in-game moves; Gobet & Simon, 1998). Slots that integrate new information help supplement short-term memory, and likely contain default values that are based on the expert’s knowledge of the frequency of occurrence of combinations of game pieces (i.e., how common a set of opening moves are in a game of chess; Gobet, 2016).

Template theory provides the necessary flexibility to account for experts’ extraordinary performance, in a way that chunking theory alone cannot. Template theory suggests that superior expert performance is supported by mental retrieval structures that can be rapidly accessed, modified to encode a game board currently in view, and/or used to make subsequent domain-specific decisions (Gobet, 2016). In conclusion, template theory is consistent with both experts’ remarkable ability to rapidly integrate and utilize new domain-specific information in short-term memory, as well as their ability to rapidly access and update domain-specific structures in long-term memory.

Taken together, chunking and template theories of expertise provide a framework for understanding the superior domain-specific perceptual abilities of chess experts (for reviews, see Bilalić, 2017; Reingold & Sheridan, 2011; Sheridan & Reingold, 2017). This framework posits that experts can rapidly access chunks and templates from long-term memory in order to process visual information as larger units rather than individual features, a process which promotes experts’ efficiency on domain-
related tasks. An effective method for testing these predictions, and for assessing perceptual expertise in general, is eye tracking. In the next section of my thesis, I will elaborate on basic eye tracking terminology, as well as describe its utility within the present research context.

**Eye Tracking**

**Basic Eye Tracking Terminology.** The eye movement record is comprised of *fixations* (i.e., periods of time during which the eye is relatively still, and visual encoding occurs) and *saccades* (i.e., periods of time during which the eye is in motion, and visual input is suppressed). During a fixation, visual information in the *fovea*, which subtends a small region of the visual field (i.e., approximately 2 degrees of visual angle), is perceived with greater clarity than visual information in the *parafovea*, which subtends a relatively larger area surrounding the fovea (i.e., approximately 5 degrees of visual angle). A critical difference between foveal and parafoveal vision is visual acuity; foveal vision is required to perceive visual information with high clarity, because visual acuity declines from foveal to parafoveal vision, and finally, to peripheral vision (see Figure 1 for an example). Despite the fact that information in the parafovea is less visually precise than information in the fovea, evidence from multiple domains suggests that the integration of visual information in the parafovea occurs to a greater extent for experts than non-experts, and may be a hallmark of perceptual expertise (for reviews, see Reingold & Sheridan, 2011; Schotter et al., 2012). Importantly, while non-experts likely have similar parafoveal vision as experts (i.e., a similar level of visual acuity), experts are likely better able to assign meaning to stimuli in the parafovea, and to perceptually group stimulus features together, using the cognitive chunks and templates developed over decades of experience.

An eye-movement parameter related to this topic is the *perceptual span*, or the amount of information that is processed in a given fixation (for a review, see Rayner, 1998). Evidence from multiple domains of expertise supports the idea that experts process domain-specific visual information in terms
of larger groups or patterns, rather than in terms of their individual features (Bilalić, 2017; Krupinski et al., 2006; Reingold et al., 2001; Richler & Gauthier, 2014; Sheridan & Reingold, 2017). This evidence suggests that experts achieve increased perceptual grouping through the use of their parafoveal vision, which results in a larger *perceptual span* for experts relative to non-experts (e.g., Reingold et al., 2001) – the size of which can be measured with the moving-window paradigm (McConkie & Rayner, 1975). The *moving-window paradigm* involves restricting participants’ gaze to a pre-defined window – everything outside of the boundaries of this invisible “window” is either blurred, or obscured from sight by a visual mask. Importantly, the window is calibrated to the movement of a participant’s eye, such that the participant can only clearly see what is directly in their foveal vision. The size of this window can be manipulated (e.g., across different blocks of an experiment), to identify the smallest possible window size for which further increases in window size produce no significant changes in performance (McConkie & Rayner, 1975; for a related discussion, see Rayner & Pollatsek, 1997).

For example, using a change detection task, Reingold et al. (2001) observed that chess experts were able to visually locate a chess piece that was rapidly changing in identity faster than non-experts, using a chess-related variant of the original flicker paradigm (Rensink et al., 1997). Implementing this technique in conjunction with the moving-window paradigm, Reingold et al. (2001) reported that experts processed more information than non-experts in a given fixation, as indicated by a larger *perceptual span*. These findings are consistent with the hypothesis that *perceptual grouping* facilitates experts’ change detection, because a larger perceptual span indicates the processing of domain-specific stimuli in terms of larger patterns.

**Benefits of Eye Tracking Methodology.** Eye tracking is particularly useful for testing the predictions of chunking and template theories. This is because the predictions of this theoretical framework involve the duration and frequency of fixations, as well as the length of the saccades. Fewer
and shorter fixations, as well as more-widely spaced fixations, are hallmarks of perceptual expertise within the eye movement record.

Additionally, eye tracking provides a highly granular continuous measure of ongoing cognition during a task. It also exists independently of self-report, which is inherently biased and less objective. Eye tracking provides an index of the time course of cognition, making it an ideal methodology for exploring nuanced differences between levels of perceptual expertise, and the theoretical frameworks that are supported by the related expertise effects. Given that eye movements are also an integral part of the task, they provide information about ongoing cognitive processing without requiring an additional response by the participant.

**Eye Tracking and Holistic Processing Perspectives.** Previously, I described the utility of eye tracking methodology for testing predictions of chunking and template theories. Importantly, expertise advantages are not uniquely predicted by chunking and template theories. Other theoretical frameworks, such as *holistic processing* perspectives (for a review, see Sheridan & Reingold, 2017) also provide a framework for the interpretation of data in domains such as medical expertise and face perception (Bilalić, 2017; Gauthier et al., 2000; Richler & Gauthier, 2014; Sheridan & Reingold, 2017; Torralba et al., 2006; Wolfe et al., 2011). *Holistic processing* refers to the way in which experts are able to rapidly obtain a “global” visual perception of a domain-specific stimulus. For example, expert radiologists can obtain the visual “gist” of a specific image, such as a chest X-ray (Drew et al., 2013). In this way, medical experts can use prior knowledge of their domain to derive meaning from this global impression, and determine the presence or absence of an abnormality in the image, such as a cancerous nodule. Experts then rapidly move their eyes to fixate the relevant region(s) of the image. This is the main premise of the *global-focal search model* of holistic processing (Nodine & Kundel, 1987; for a review, see Sheridan & Reingold, 2017).
Interestingly, and similar to chunking and template theories, the *global-focal search model* assumes that experts rely on large-scale mental templates (or schemas) to support performance. As such, these two theoretical frameworks are not mutually exclusive; global perception is facilitated when experts are viewing groupings of domain-specific familiar visual features. Because my dissertation research cannot directly address whether either chunking and template theories, or holistic processing theories, are uniquely contributing to the observed results, an elaborate discussion of holistic processing is outside of the scope of this report. More research is needed to dissociate the independent contributions and predictions of both chunking processes and global perception processes within domains like music reading expertise.\(^4\)

In the next section of my thesis, I will elaborate on the work that sets the precedent for my dissertation research. I will first discuss several important perceptual expertise effects observed in the chess domain. Next, I will provide a discussion of the similarities between the domains of chess and music reading, detailing the rationale for exploring whether support for a chunking and template framework could be observed in the domain of music. I will conclude this section with a review of important perceptual expertise effects in the domain of music reading.

**The Domain of Chess**

Explorations of perceptual expertise effects in the chess domain have played a critical role in shaping research on visual perception and skill acquisition over the past few decades. Simon and Chase (1973) highlight the fundamental role of chess research in understanding human knowledge acquisition with a compelling analogy; much as *Drosophila* (the fruit fly) is viewed as a model organism in the study of genetics, chess can be viewed as a model organism in the study of perceptual expertise.

---

\(^4\) See “Limitations and Future Research Directions” for further discussion.
One of the most substantial and unique methodological advantages of the domain of chess is its nationally recognized system for ranking players based on their level of expertise (i.e., the Elo rating; Elo, 1978). The Elo rating system assigns players a numerical rating (range: 0 to 3000) that changes depending on game outcomes. Elo ratings are used for determining tournament eligibility, pairing competitors, and predicting game outcomes. For the cognitive scientist, the Elo rating system provides an elegant way to operationalize expertise. Given the fundamental nature of the issue of definition to the field of expertise research at large, the importance of the Elo system cannot be overstated.5

Additional aspects of the game of chess further highlight its role as a “model organism” for studying perceptual expertise. The game of chess follows a defined set of rules and is played with a finite set of 32 game pieces, 16 of which are controlled by an individual player. Within this set of 16, there exist six unique piece types (pawn, bishop, knight, rook, king, and queen) that interact within a finite set of legal moves. The relatively small set size and the predictability of game piece movement makes the domain of chess ideal for timed behavioral experiments. Additionally, unlike domains such as sports expertise, it is more feasible (and more ecologically valid) to study chess expertise in an controlled laboratory setting.

In summary, the domain of chess is ideal for studying and understanding perceptual expertise effects because of the algorithmic way that chess expertise is defined, the finite number of pieces and defined patterns of movement of these pieces, and the relative ease of access to chess experts as participants in comparison to other experts, like radiologists, who likely have differing constraints on their time that make recruitment for a behavioral experiment potentially challenging.

5 For further discussion, see Defining Expertise in “General Discussion”.

10
Perceptual Expertise Effects in the Domain of Chess

Seminal studies in the chess domain revealed that experts can remember and reconstruct a large number of game piece configurations from briefly presented chessboards (de Groot, 1946). However, this ability is greatly diminished if experts are shown randomized chessboards instead of pieces in legal configurations (Chase & Simon, 1973a, 1973b; Gobet & Simon, 1996, 2000). As described earlier, these findings serve as the foundation for the theoretical framework underlying our current understanding of chess expertise (chunking and template theories). These findings also provide compelling support for the conclusion that the perceptual abilities of chess experts can be attributed to their extensive knowledge of domain-specific visual patterns, and not a general memory or perceptual advantage (Charness et al., 2005; Reingold & Sheridan, 2011). The fact that experts’ memory for chess pieces is not uniformly superior for both legal and random configurations implies that an expert advantage cannot be attributed to superior general memory (relative to non-experts). In their review, Reingold and Sheridan (2011) detail the pivotal role of chess research in initiating a shift away from viewing chess expertise as a result of innate talent or skill, and toward viewing chess expertise as a result of domain-specific knowledge.

Reingold et al. (2001) provided compelling support for the domain-specificity of chess expertise, using a change detection task in conjunction with a moving window technique (as described earlier). Participants engaged in a check detection task, indicating whether a King was in check (i.e., whether the player is in danger of losing the game) by surrounding pieces on a simplified 3x3 chess board. Reingold et al. (2001) reported that experts were better at this task than non-experts, as indicated by more efficient eye movements (i.e., fewer fixations and longer saccades). Reingold et al. (2001) also documented a larger visual span by experts during a domain-related change detection task. A large visual span is heavily supported by the experts’ parafoveal vision (see Figure 1); this manifests in the eye movement record as few saccades during a given trial, and fixations in between (rather than on)
individual pieces. Experts who are able to encode meaningful information from the area around their point of fixation possess a larger visual span, and can complete trials more efficiently than those with smaller visual spans.

In order to demonstrate that chess experts’ encoding advantage is domain-specific, and not the result of an enhanced general perceptual ability, Reingold et al. (2001) manipulated the familiarity of the notation used to represent the chess pieces (i.e., symbols, like a King piece vs. letters, like a “K” to represent a King). Importantly, while the symbols are familiar to the experts and the letters are not, both conditions provide the same information about the identity of a given chess piece. As such, to the extent that experts’ encoding advantage is domain-specific, superior performance should be observed for the familiar condition only. In support of this prediction, Reingold et al. (2001) observed a significant reduction in the efficiency of experts’ eye movements in the letter condition only, demonstrating that this encoding advantage is due (at least in part) to experts’ practice with chess symbols, and not a general memory superiority.

More contemporary work using both eye tracking and neuroimaging techniques demonstrates that chess experts’ perceptual advantages stem from their domain-specific knowledge of visual chunks (Bilalić, 2017; Bilalić et al., 2010). Chess experts in Bilalić and colleagues’ (2010) experiment completed a domain-specific visual search task with more efficient eye movements than non-experts; experts rapidly fixated relevant regions of the stimulus, while non-experts fixated both relevant and irrelevant information. As in Chase and Simon (1973a) and Reingold et al. (2001), experts’ efficiency was greatly diminished (but not eliminated) when presented with random (and not legal) chess configurations, relative to non-experts.

In this section, I have described the methodological strengths of the domain of chess, as well as the seminal findings in this domain that motivated my thesis research. Throughout the thesis, I have also
highlighted the ways in which chunking and template theories can robustly account for these findings in the chess domain. To understand the connection between this discussion and my current work in the music reading domain, I will next elaborate on the similarities between the domains of chess and music reading. This discussion sets the precedent for considering an extension of chunking and template theories to a novel domain.

**The Domains of Chess and Music Reading**

Both the domains of chess and music reading involve complex meaningful visual patterns that are comprised of smaller individual components with legal and random configurations (Madell & Hébert, 2008; Reingold et al., 2001; Reingold & Sheridan, 2011; Sheridan et al., 2020). The individual visual features of a music score can be grouped into larger visual patterns that become familiar to experts over the course of many hours of practice. For example, individual chess pieces on a board and music notes within a score can form patterns that could be described as “chunks” (e.g., “pins” or “forks” in chess, and “arpeggios” and “chords” in music). These patterns can, in turn, comprise even larger patterns that could be described as “templates” (e.g., entire opening board configurations in chess, and melodies in music) (see Figure 1 for an example). Because both domains involve perceptual grouping, it would be interesting to explore the extent to which chunking and template theories could be extended to the domain of music.

In addition to the similarities between the domains of chess and music, there are multiple strengths of music reading as a domain of expertise that make it ideal for studying the boundary conditions of perceptual expertise effects. The symbols used in music reading studies (i.e., music notes, expressive markings, etc.) are more numerous, complex, and variable than the symbols used in studies of chess expertise (i.e., a finite set of 32 chess pieces per game). This makes music reading a promising domain for testing the predictions of chunking and template theories. This diversity permits more
granular levels of comparison between conditions, which results in a stronger manipulation of variables like visual complexity than are possible in domains, like chess or medicine. Music reading expertise is also uniquely multisensory, often involving simultaneous music performance (auditory and motor) and music reading (visual). Methodologically speaking, it is relatively easier to secure large samples of expert musicians than it is to secure comparably sized samples of chess experts. For these reasons, music reading is an ideal domain for studying the boundary conditions of perceptual expertise effects.

Compared to domains of expertise like chess and medicine, relatively less work has explored visual expertise in the domain of music (for reviews, see: Bilalić, 2017; Madell & Hébert, 2008; Sheridan et al., 2020). This may be at least partially due to the fact that there has historically been a much greater focus on the auditory (rather than the visual) component of music expertise (for a review, see Madell & Hébert, 2008). Highlighting this discrepancy, reviews of the music reading and eye tracking literature have indicated the need for future work focused on developing a more comprehensive understanding of eye movement behavior during music reading (Madell & Hébert, 2008; Puurtinen, 2018). Formally exploring the extension of theories of visual expertise (i.e., chunking and template) to the domain of music reading expertise will help to advance the field’s understanding of visual expertise as a whole.

**Perceptual Expertise Effects in the Domain of Music Reading**

My dissertation builds on the prior music reading expertise and eye tracking literature (for reviews, see: Madell & Hébert, 2008; Puurtinen, 2018; Sheridan et al., 2020). Within this literature, it is generally found that expert music readers make fewer fixations than non-experts, and longer saccades, during a music reading task (Sheridan et al., 2020). This pattern of eye movement behavior is considered to be a hallmark of perceptual expertise, and a display of expert efficiency. Experts rely on this efficiency in order to complete domain-related tasks more quickly and accurately than non-experts (for a review, see Reingold & Sheridan, 2011).
In many fields of expertise, experts display shorter fixation durations during domain-related tasks (chess: Reingold & Sheridan, 2011; medicine: Sheridan & Reingold, 2017; reading: Rayner, 2009). However, within the music reading domain, there are mixed results about the effects of expertise on fixation duration. Sometimes, expert music readers exhibit shorter fixations than non-musicians, which is likely reflecting experts’ enhanced speed of processing (Goolsby, 1994a, 1994b; Penttinen et al., 2013; Truitt et al., 1997; Waters & Underwood, 1998; Waters et al., 1997). But other times, the opposite pattern of results is observed, where experts exhibit longer fixations than non-musicians (Gilman & Underwood, 2003; Maturi & Sheridan, 2020). This result may reflect the experts’ ability to both encode more information in a single glance than non-experts, as well as utilize more information present in their parafoveal vision than non-experts (for a review of these effects in the chess domain, see Reingold & Sheridan, 2011).

This latter pattern of results directly supports the conclusion that experts are engaging in chunking behavior; longer fixations suggest that experts are spending more time than non-experts encoding domain-related patterns. More specifically, the domain of music may support chunking behavior among experts because the visual symbols themselves delineate which features are perceptually grouped together in a chunk. Unlike chess, music scores contain a large number of visual cues that could influence the likelihood that individual elements would be perceptually grouped and perceived together. For example, the presence of a “beam” in written music (i.e., a line connecting two notes) is a visual cue that could facilitate chunking during music reading (Madell and Hébert, 2008).

Experts may also spend longer looking in a certain location in order to engage their parafoveal vision. Eye movements are required to process visual information because this information is viewed with the greatest clarity at the center of a fixation, or in foveal vision. Foveal vision subtends a small region of the visual field (approximately 1 degree of visual angle from the point of fixation). Additionally, the importance of parafoveal vision (i.e., the part of the visual field surrounding the foveal region; up to
5 degrees of visual angle from the point of fixation) to perceptual expertise has been robustly
documented in multiple domains (for reviews, see Reingold & Sheridan, 2011; Sheridan et al., 2020).
Experts are able to utilize their parafoveal vision to encode more visual information in a single glance
than can non-experts. This is yet another hallmark of perceptual expertise and efficiency of eye
movements, and again supports experts’ faster and more accurate task performance.

Within the music reading expertise and eye tracking literature, there is evidence to suggest that
chunking may be occurring among experts. For example, some studies have observed that expert
musicians will fixate on blank spaces within a music score during a task (Gilman & Underwood, 2003;
Goolsby, 1994a, 1994b; Madell and Hébert, 2008; Truitt et al., 1997). These fixations occur relatively
frequently, presumably reflecting the importance of this type of fixation; two studies reported that the
percentage of expert fixations on blank space in a music score was roughly equivalent to the percentage
of fixations on music notes (Gilman & Underwood, 2003; Truitt et al., 1997). This behavior may initially
seem counterproductive, as the purpose of fixation is to bring visual information into clearer focus for
encoding purposes. However, this finding has a parallel in the chess expertise literature, where chess
experts have also been observed fixating on the blank space in between individual chess pieces
(Reingold & Sheridan, 2011; Sheridan et al., 2020).

Why might fixation on blank space be advantageous for an expert? To the extent that experts
are relying on parafoveal vision to process domain-related stimuli, they would be able to encode more
visual information than non-experts who do not (or cannot) rely on parafoveal vision to this same
extent. Blank spaces of fixation are likely surrounded by meaningful visual symbols, and thus it may be
advantageous to utilize parafoveal vision to encode a greater amount of meaningful information than
would be possible with foveal vision alone. Fixation on blank space likely allows experts to mentally
chunk the relevant information in their parafoveal vision to a greater extent than would be possible with
foveal vision alone. In this way, the use of parafoveal vision supports chunking behavior among experts.
Relatedly, there is growing evidence in the music reading expertise literature that expert musicians are engaging in more parafoveal processing than are non-musicians (for a review, see Sheridan et al., 2020). This is evidenced by experts’ larger *eye-hand span* (EHS) than non-musicians’. It is often the case that during performance, a musician will be looking ahead of where they are actually performing within the music score. This helps the performer anticipate and execute stylistic or expressive markings within the score. The *eye-hand span* can be understood as the distance (in absolute time; number of musical beats; etc.) between where the eye is fixating and the location of the performer’s hands (for a review, see Perra et al., 2021).

Thus far, I have discussed the most important findings in the perceptual expertise literature from the chess domain, as well as findings from the music reading domain. I described some of the similarities between perceptual expertise effects in both of these domains. Importantly, I also discussed Reingold et al. (2001), which motivated the core methodology of my dissertation by setting a precedent for the utility of change detection paradigms for exploring differences between levels of expertise. However, consistent expertise advantages are not always observed within the change detection literature. In the next section of my thesis, I will discuss these inconsistencies and describe how the two experiments in my dissertation help to clarify the boundary conditions under which chunking occurs. This helps to fill a gap in the literature by assessing the generalizability of chunking and template theories to a novel domain of expertise.

The Change Detection Literature

Broadly speaking, my dissertation research investigates the mechanisms by which long-term memory representations influence perceptual encoding, which is a long-standing question in both the change detection literature (Rensink et al., 1997; Simons & Rensink, 2005) and the visual expertise literature (Reingold & Sheridan, 2011). The ability to rapidly detect a change in an image across time is
surprisingly challenging (Simons & Armbinder, 2005), requires both attention and fixation (Hollingworth et al., 2001; Rensink et al., 1997), and has important real-world implications, such as safety while driving a vehicle (Galpin et al., 2009; Groff & Chaparro, 2003) and eyewitness testimony (Laney & Loftus, 2009). The failure to detect such a change is called change blindness or inattentional blindness.

Experts are not immune to these striking failures to notice cues that would otherwise be very salient (Simons & Rensink, 2005). In an especially compelling demonstration of expert change blindness, Drew, Võ, and Wolfe (2013) observed that 83% of expert radiologists failed to notice a gorilla that was 48 times the size of an average target during a domain-specific visual search task (i.e., searching for a cancerous nodule in a radiograph of a lung). In addition, the majority of the experts who failed to notice the gorilla had looked directly at the gorilla’s location, as revealed by eye tracking measures. This result suggests that attention and fixation do not always necessarily co-occur. This study in particular highlights the utility of eye tracking methodology in exploring questions related to change blindness. One reason for a failure to notice a salient cue could be the lack of fixation on that cue; eye tracking helps reveal where and when a participant is looking, to confirm whether a cue was indeed fixated.

However, expert performance during change detection tasks is not uniformly poor; the overall literature on expertise effects during change detection is inconsistent. Sometimes experts show performance advantages relative to non-experts, and sometimes they do not. In studies with expert football players (Werner & Theis, 2000), expert chess players (Reingold et al., 2001), and expert drivers (Groff & Chaparro, 2003; Zhao et al., 2014), perceptual expertise effects were observed as experts outperformed non-experts during a change detection task.

However, in other studies with expert soccer players (Cañal-Bruland et al., 2011) and expert video game players (Gobet et al., 2014), experts and non-experts did not differ in reaction time during a domain-related change detection task. The effect of driving expertise on change detection seems to be
especially context-dependent (Zhao et al., 2014). When the change detection task is embedded within another task that simulates real-world driving (i.e., checking a rear-view mirror for hazards and monitoring the distance between themselves and a vehicle in front of them), better drivers outperform poorer drivers (Crundall, 2009; Koustanaï et al., 2012). However, when completing a similar change detection task in isolation, expertise effects disappear (Galpin et al., 2009; Mueller & Trick, 2012).

These inconsistencies within the change detection literature merit further investigation. In the next section of my thesis, I will introduce the novel change detection paradigm that I adapted from Reingold et al. (2001) to explore the boundary conditions of expertise effects on change detection in the music reading domain. I will also describe the primary goals of my dissertation, as well as my specific predictions for my two experiments.

The Dissertation

Empirical support for the idea that experts are engaging in chunking behavior has been observed robustly in the domain of chess expertise (Reingold & Sheridan, 2011). As described earlier, patterns of eye movement behavior that reflect efficiency in experts (i.e., fewer fixations and longer saccades relative to non-experts) also support the conclusion that experts are engaging in chunking behavior. Another methodological strength of exploring the domain of music reading expertise is the large number of tasks that are available to explore expertise effects among musicians. These tasks vary greatly in terms of difficulty and modality (e.g., music-reading tasks, such as sight singing, sight reading, and silent sight reading; music-performance tasks, such as silent performance; memory tasks, such as music-related visual search and pattern matching tasks). The degree to which chunking is occurring in response to variable task demands in music reading remains an open question. To further understand the boundary conditions of chunking and template theories, my dissertation explores this question with one specific methodology, namely a change detection task.
The primary goal of my dissertation is to clarify the boundary conditions of expertise effects on change detection. Two experiments focus on exploring music reading experts’ domain-related change detection ability. This work helps clarify the conditions under which chunking occurs and impacts expert change detection (for further discussion, see Sheridan & Kleinsmith, 2021). In this way, my work fills a gap in the literature and assesses the generalizability of chunking and template theories to a novel domain of expertise.

In my dissertation, I developed a novel variant of the change detection paradigm, which was motivated by the chess-related variant of the flicker paradigm (Rensink et al., 1997) that was introduced by Reingold et al. (2001). In Reingold et al.’s (2001) study, chess experts and non-experts were required to detect (i.e., fixate) a chess piece that was rapidly changing in identity. Reingold et al. (2001) observed that experts completed trials containing legal game positions with larger visual spans than non-experts, meaning that experts were able to encode more information in a single glance. Critically, larger visual spans by experts were not observed for trials containing randomized game positions, suggesting that these expertise advantages only manifest when the to-be-remembered item is perceptually specific to the context and format in which it was learned or practiced. While these findings are consistent with chunking and template theories’ prediction that perceptual expertise is domain-specific, more work is needed to explore the boundary conditions of this theoretical framework within the domain of music reading expertise.

By extending this paradigm to the music reading domain, my goal is to explore the boundary conditions of expertise effects on change detection. As mentioned previously, this literature is not consistent within itself, sometimes reporting expertise advantages during change detection, and sometimes reporting no expertise effects during change detection. By introducing a change-detection version of the flicker paradigm within music, this work contributes to the growing number of tasks that can be used to explore music reading expertise. Another strength of this task is its accessibility to true
non-experts (or, naïve observers); a participant does not need to know how to read music to locate a changing target. This permitted a strong manipulation of expertise, reflected in the large and equal sample sizes in each of my experiments ($N = 60; 30$ experts and $30$ non-musicians). Donovan and Litchfield (2013) suggest that a gap exists in the visual expertise literature concerning data from true non-experts. It is often the case that participants in the “novice” group in expertise studies have at least some understanding of their domain, which is required in order to complete the task. Because of this, comparatively less is understood about the performance of naïve observers, which underscores the importance of understanding the behavior of these participants.

To summarize thus far, I have described how chunking and template theories and their assumptions are consistent with perceptual expertise effects shown in the domain of chess (Reingold et al., 2001; Reingold & Sheridan, 2011; Sheridan et al., 2020). For example, experts tend to exhibit fewer fixations and longer saccades relative to non-experts during a domain-related task. To explore the boundary conditions of perceptual expertise effects, and to understand the generalizability of these effects, I tested assumptions of chunking and template theories in my dissertation, with a focus on the domain of music reading. I described the strengths of music reading as a domain for studying expertise effects. A wide variety of tasks that provide different ways and modalities of studying music reading expertise, ease of access to large and robust samples of expert music readers, and the rich complexity of the stimuli themselves permitting strong comparisons between groups and conditions, combine to make the domain of music reading expertise ideal for studying the boundary conditions of perceptual expertise effects. The task that I developed adds to the large number of tasks available for studying music reading expertise, permits a comparison across domains to the chess literature where this task was used previously, and is accessible to a true novice group.

My dissertation work addresses several gaps within the music cognition and perceptual expertise literatures. In this work, I am exploring the idea that chunking and template theories can be
applied to the novel domain of music reading expertise, in addition to exploring the boundary conditions of expertise effects on change detection. This work also contributes to the extant literature that examined eye movements during music reading, which is relatively small as compared to the larger body of work that has examined eye movements during text reading (for reviews, see Rayner, 1998, 2009). Previous reviews of the music reading and eye tracking literature highlight the need for future work aimed at further understanding how changes in the physical/written components of a music score impact eye movement behavior (Madell & Hébert, 2008), and my dissertation directly contributes to this effort. Relatedly, Fink et al. (2018) notes that “Many authors assert the necessity of agreeing upon ... dependent measures to increase comparability within the music research community and beyond to other research domains” (p. 3). By developing a novel variant of a paradigm that has reliably revealed expertise effects on change detection in the domain of chess, my dissertation work provides a foundation for a direct comparison between the domains of chess and music. Finally, and more broadly, my dissertation contributes to ongoing efforts to develop a comprehensive and unified theory of visual expertise (e.g., Brams et al., 2019).

To assess the extent to which visual expertise effects occur in the domain of music, I introduced a novel music-related variant of the flicker paradigm (Rensink et al., 1997). Specifically, I monitored the eye movements of 60 expert musicians (30 per experiment; with at least 10 years of music reading experience) and 60 non-musicians (30 per experiment; who could not read music) while they located a single music note or note stack that was rapidly appearing and disappearing within a larger music score. Across two experiments, I explored the impact of manipulating visual elements of the music scores on experts’ change detection ability. In Experiment 1, I manipulated visual complexity to test the prediction that experts’ domain-specific knowledge permits them to compensate for increases in visual complexity. In Experiment 2, I manipulated the visual configuration of the music score by contrasting upright (i.e., typically oriented and familiar) scores with those that were rotated by 90° (i.e., unfamiliar) to test the
prediction that visual expertise is perceptually specific. To foreshadow my findings, I observed support for the prediction that chunking permits experts to compensate for increases in task complexity in ways that non-experts cannot, as well as support for the prediction that music reading expertise (like chess expertise) is remarkably perceptually specific.

**Experiment 1**

In Experiment 1, I introduced a novel music-related variant of the flicker paradigm (Rensink et al., 1997), which was inspired by the chess-related variant introduced by Reingold et al. (2001). Reingold et al. (2001) observed that compared with less skilled players, experts were faster at detecting chess pieces that were rapidly changing in identity, and processed more information in a given fixation, as indexed by a larger visual/perceptual span (for a related discussion, see Reingold & Sheridan, 2011). Critically, experts’ advantages in change detection only occurred for legal (and not randomized) game positions, supporting Reingold et al.’s (2001) conclusion that chess experts’ superior performance on this task relies on domain-specific perceptual encoding advantages, and not general cognitive advantages.

Specifically, Experiment 1 was designed to assess whether these patterns of results from the chess domain could generalize to the music reading domain. In the subsequent paragraphs, I will describe the specific predictions of Experiment 1, as well as Experiment 1’s Method, Results, and Discussion.

**Experiment 1 Predictions**

Experiment 1 was designed to test the specific prediction of chunking and template theories that experts’ domain-specific knowledge permits them to compensate for increases in task complexity in ways that non-experts cannot. Experts’ ability to rapidly recognize complex visual patterns is a key component of chess expertise (e.g., Sheridan & Reingold, 2017). Sheridan & Reingold (2017) used a task
that required expert and non-expert participants to rapidly conduct a visual search for a single target within an array of lures. The target was the only board configuration that allowed the knight piece to reach a target square in three moves. Eye movement evidence supporting the conclusion that experts can rapidly differentiate complex visual patterns was observed: experts (but not non-experts) showed longer first-fixation durations on the target board, relative to a lure that was difficult to reject. Here, experts’ longer first fixations permitted remarkably rapid differentiation between complex visual patterns (i.e., at first glance) in ways that non-experts could not. Indeed, “this difference was solely driven by [expert] trials that elicited a single-dwell on the target board” (p. 8).

As discussed earlier, there is ample reason to consider the similarities between chess and music, as well as consider the extension of theoretical frameworks supported in the chess domain to the music reading domain. Experiment 1 permits a test of the broader hypothesis that chunking facilitates change detection, as well as explores whether this mechanism extends beyond the chess domain. Music is similar to chess because music scores contain individual features (e.g., music notes) that can be grouped into larger meaningful patterns (e.g., “chunks”, such as chords and arpeggios). An important advantage of the domain of music is that written music varies dramatically in visual complexity (Figure 1); because of this, I was able to implement a stronger manipulation of visual complexity than would be possible in the domain of chess. This makes music an ideal domain for studying the boundary conditions of expertise effects on change detection.

Complex music scores contain more visual features than simple scores, thus increasing participants’ visual working memory load during trials containing complex scores. Chunking is known to reduce the burden of an increased working memory load (Thalmann et al., 2019), and prior research suggests that change detection can provide an index of visual working memory load (Pailian & Halberda, 2015). For these reasons, to the extent that experts are using chunking during change detection, I predict that experts will be less affected by increases in visual complexity than non-musicians. In
contrast, no evidence for chunking would be observed if experts and non-musicians are equally affected by the visual complexity manipulation.

I tested the predictions of Experiment 1 with eye tracking measures (saccade amplitude and fixation duration) and traditional behavioral measures (accuracy and reaction time (RT)). Building on the prior work in the chess domain (for a review, see Sheridan & Reingold, 2011), I predicted that expert musicians would show faster RTs and higher accuracy, relative to non-musicians. This is because chunks can permit experts to process more information in a given fixation, which supports efficient performance (Maturi & Sheridan, 2020; Reingold & Sheridan, 2011; Sheridan et al., 2020). To this end, I predicted that expert musicians would show fewer number of fixations and larger saccade amplitudes than non-musicians; processing more information in a given fixation can permit more widely-spaced fixations, without loss of information in experts’ working memory. Finally, I did not have an a priori prediction for the fixation duration measure; results have been mixed about fixation duration changes as a function of expertise (for a review, see Reingold & Sheridan, 2011). Experts may make longer fixations, during which time they can more deeply encode visual information than non-experts, but experts may also make shorter fixations, which is a hallmark of their efficiency with processing domain-specific stimuli (Reingold & Sheridan, 2011).

Critically, I predicted that expertise effects on saccade amplitude, accuracy, and RT would be magnified in the complex relative to the simple condition. As described above, “chunks” can facilitate experts’ encoding during a domain-specific task. Similarly, experts’ use of “chunks” should also help compensate for the increase in visual working memory load during complex (relative to simple) trials.
Method

Participants

A total of 60 participants were recruited for Experiment 1. The expert participants were recruited from the University at Albany, SUNY, and the surrounding community (n = 30; 13 females), and the non-musician participants were recruited from the University at Albany, SUNY’s research participant pool (n = 30; 16 females). As compensation, experts received $10/hour and non-musicians received course credit. The experts had a minimum of 10 years of music reading experience (M = 13.8 years; range = 10–39 years), including music theory training and music performance of at least one instrument. The non-musicians were not able to read music. The expert’s mean age was 25 (median = 20; range = 18–55) and the non-musician’s mean age was 19 (median = 18; range = 18–21).

The experts in Experiment 1 played a wide range of instruments: piano (11), guitar (10), bass guitar (6), violin (5), trombone (5), trumpet (4), trombone (4), alto saxophone (3), flute (3), clarinet (3), viola (3), ukulele (3), cello (2), French horn (2), organ (2), tuba (1), tenor saxophone (1), piccolo (1), bassoon (1), oboe (1), and ocarina (1). 90% of experts played more than one instrument.

Experts who exclusively and/or primarily played percussion instruments were excluded, as most percussion sheet music does not contain melodic contour in the way that brass or woodwind sheet music does. All experts were actively engaged in playing a melodic instrument and reading sheet music at the time of the experiment, and nearly all experts (97%) had no lapses in their music training (i.e., had not stopped practicing and reading music since they began). All experts had studied music formally in an educational setting, and most (93%) reported having at least “advanced” knowledge of music theory.

---

6 In Experiment 1, two expert-level participants were identified with recruitment efforts within the research pool, and as such, received course credit instead of monetary compensation for their participation.

7 Contour is known to be an important factor in long-term memory for music (Bartlett & Dowling, 1980; Dowling, 1978; Kleinsmith & Neill, 2018), and is an equally important component of music expertise (Talamini et al., 2022).
Most experts (93%) also studied music informally, by watching videos, completing composition and performance requests for others, or practicing “by ear” (i.e., listening to a piece of music and attempting to replicate it without the aid of sheet music). See Appendix A for the complete demographic questionnaire.

Materials and design

The experimental materials were composed of 100 four-line excerpts from lesser-known Classical era music scores drawn from a variety of instruments (piano, French horn, flute, clarinet, and violin). All stimuli were roughly rectangular in shape, with slight natural variations in the absolute size of the music score.\(^8\) Fifty of these images were categorized as “visually simple” (e.g., Figure 3, Panel A) and 50 were categorized as “visually complex” (e.g., Figure 3, Panel B) based on complexity ratings collected from five independent raters. These raters were instructed to define complexity as the “amount of ink on the page” on a Likert scale ranging from 1 (not at all complex) to 5 (very complex). To create a strong complexity manipulation, I only included music scores that received extreme average complexity ratings as experimental trials (i.e., < 2 for simple and > 4 for complex). An additional four music scores (two complex and two simple) were included as practice trials at the beginning of the study. Each participant saw a given music score only once.

To implement the flicker paradigm (Rensink et al., 1997), altered versions of each music score were also created. The altered version was identical to the original music score except that one music note or one chord (a chord is a vertical stack of music notes) was removed from the image. The target locations were randomly selected to ensure that the location of the change would not be predictable from trial-to-trial. Specifically, each musical score was first divided into four “quadrants”, with each

---

8 Simple condition: width range: 13.0° to 14.1°; height range: 7.2° to 8.0°
Complex condition: width range: 13.3° to 14.2°; height range: 7.0° to 8.0°
quadrant containing approximately five bars of music. For each image, the target was assigned randomly to one of these quadrants with the constraint that across all images there was an equal chance (i.e., 25%) of the target appearing in each of the four quadrants. After assigning the target to a quadrant, I then randomly assigned the target to a specific bar within the quadrant (each bar within a quadrant had an equal chance of being selected). Finally, I randomly selected the specific note or chord within the bar that would serve as the target for a given image (each note within a bar had an equal chance of being selected).

I manipulated visual complexity within-subjects, and used a 2 (Complexity: Visually Complex or Visually Simple) × 2 (Expertise: Experts or Non-musicians) design.

**Apparatus**

Eye movements were monitored using the SR Research EyeLink 1000 Plus system, which has high spatial resolution (gaze-position error was less than 0.5° after calibration) and high temporal resolution (the sampling rate was 1000 Hz). Viewing was binocular, but only the right eye was monitored. A chin and forehead rest were used to minimize head movements. The music scores were displayed on a 24-inch BENQ XL2420-B computer screen with a resolution of 1,920 × 1,080 pixels and a refresh rate of 144 Hz. Saccades were detected using the standard parser settings for the SR Research EyeLink 1000 Plus system. Specifically, the velocity threshold was 30°/s and the saccade acceleration threshold was 8,000°/s². The screen was 70 cm away from the participant. The music notation in the images was shown in black on a white background, and the music score was always displayed in the center of the screen with a gray border around the edges of the screen. There were slight variations in the size of the music scores. A gamepad was used to collect responses.
Procedure

At the start of each trial, participants were instructed to look at a fixation cross that was located in the center of the screen. Their fixation on this cross initiated the trial. Participants then completed a flicker paradigm task (Rensink et al., 1997). As illustrated by Figure 3, Panel C, the original and altered versions of the music score were briefly presented (for 700 ms each) and continuously alternated with a blank screen shown between each music score screen (for 100 ms). This presentation sequence continued either until a response was made or 2 minutes had elapsed. The time-out value of 2 minutes was reached on less than 1% of trials, and these trials were excluded from my analyses. When the participants located the changing note or chord, they responded by fixating on its location while pressing a button on a gamepad. Once the button was pressed, their fixation location was recorded and the trial ended. Participants completed four practice trials, followed by 100 experimental trials (i.e., 50 visually simple trials and 50 visually complex trials). The trials from the two complexity conditions were randomly intermixed. Participants were encouraged to take frequent breaks in between trials.

Results

To examine the impact of expertise and visual complexity on change detection, I will first report accuracy and reaction times (RTs), followed by three eye tracking measures (fixation duration, number of fixations, and saccade amplitude). For all measures reported below, I conducted 2 × 2 ANOVAs, with Complexity (simple, complex) as a within-subjects variable and Expertise (expert, non-musician) as a between-subjects variable. Error trials (i.e., trials in which the changing note was not accurately located) were removed from all analyses. Experiment 1 data are presented in Table 1.

RT and Accuracy

Figure 4 displays the means and standard errors for the accuracy (Panel A) and RT (Panel B) measures. The experts were faster than non-musicians ($F(1,58) = 50.59, p < .001$). Accuracy did not differ
as a function of expertise ($F(1,58)= 2.83, p = .098$). The simple trials elicited faster and more accurate responses than the complex trials (RT: $F(1,58) = 407.23, p < .001$; Accuracy: $F(1,58) = 8.82, p = .004$).

In support of my predictions, expertise effects on performance were larger for complex than simple trials, as indicated by significant interactions for both RT ($F(1,58) = 15.11, p < .001$) and accuracy ($F(1,58) = 6.52, p = .01$). Follow up t-tests revealed significant expertise effects on RT in both conditions (simple: $t(58) = 8.11$; complex: $t(58) = 5.94$, both $p$’s < .001). These t-tests also revealed significant expertise effects on accuracy in the complex ($t(58) = -2.14, p = .037$), but not in the simple condition ($t(58) = -0.02, p = .986$).

**Fixation duration and Number of fixations**

Figure 4 displays the means and standard errors for the fixation duration (Figure 4, Panel C) and the number of fixations (Figure 4, Panel D) measures. Relative to the non-musicians, the experts had shorter fixation durations ($F(1,58) = 14.23, p < .001$), and fewer numbers of fixations ($F(1,58) = 23.27, p < .001$). Relative to complex trials, the simple trials elicited shorter fixation durations ($F(1,58) = 171.48, p < .001$) and fewer numbers of fixations ($F(1,58) = 296.35, p < .001$).

More importantly, and in support of my predictions, expertise effects were larger in the complex than the simple condition for the number of fixations measure ($F(1,58) = 5.99, p = .017$); follow-up t-tests showed significant expertise effects on the number of fixations in both conditions (simple: $t(58) = 5.34$; complex: $t(58) = 4.02$, both $p$’s < .001). In contrast, the fixation duration measure did not show a significant interaction ($F(1,58) < 1, p = .946$). This reveals that experts’ performance advantage in the complex condition was reflected in changes in the number of fixations per trial, rather than changes in fixation durations.
**Saccade amplitude**

Figure 4 displays the mean and standard error for the saccade amplitude measure (Figure 4, Panel E). In support of my predictions, experts had larger saccade amplitudes than the non-musicians ($F(1,58) = 18.48, p < .001$), and saccade amplitudes were significantly larger for the simple, as compared to the complex, condition ($F(1,58) = 90.91, p < .001$). More importantly, I observed a significant expertise by complexity interaction for saccade amplitude (i.e., expertise effects on saccade amplitude were larger in the complex condition), ($F(1,58) = 5.07, p = .028$). This interaction suggests that experts are better able than non-experts to compensate for increases in task complexity, in addition to suggesting that experts have a larger visual span than non-experts.

**Experiment 1 Discussion**

Experiment 1 tested the predictions of chunking and template theories that experts’ domain-specific knowledge permits them to compensate for increases in visual complexity during a domain-specific task. In support of this prediction, I observed significant expertise effects on RT, accuracy, and number of fixations, as a function of complexity. Experts completed all trials faster than non-musicians, and were more accurate than non-musicians on complex trials (there was no difference in accuracy between groups for simple trials).

As described earlier, chunking mechanisms would be implicated during change detection if experts were less affected by increases in visual complexity than non-musicians. In this way, experts’ ability to complete complex trials with fewer fixations than non-musicians supports the conclusion that chunking is occurring, and facilitating change detection among experts. Specifically, increases in visual complexity magnified the effect of expertise on fixation count. Significantly larger saccade amplitudes for experts than non-musicians (Figure 4, Panel E) suggests that there might be a larger visual span for music reading experts relative to non-musicians for this task (for a review, see Sheridan & Reingold,
Future work could further explore the individual differences in visual span during the current task.

Taken together, these results provide support for the conclusion that experts are able to rely on mental chunks to facilitate superior performance relative to non-musicians, and to compensate for increases in visual complexity (i.e., which tax visual working memory load) during a domain-specific task. In line with the predictions of chunking and template theories, experts may be better able to engage in chunking techniques in the complex condition of Experiment 1, as this condition is characterized by significantly more “ink on the page”, as compared to the simple condition. As a result, experts’ superior performance in the complex condition relative to non-musicians may be due, at least in part, to the fact that there is simply more visual information that can be chunked in the complex condition.

Experiment 2

The perceptual specificity of memory representations has been a long-standing issue in the field of visual expertise research (Bilalić, 2017; Reingold, 2002). A hallmark of visual expertise is the ability to rapidly and efficiently encode domain-specific visual patterns (Chase & Simon, 1973a; Gale et al., 2000; Kundel & Nodine, 1975; Nodine & Krupinski, 1998; Nodine & Kundel, 1999; Reingold et al., 2001; Reingold & Sheridan, 2011). However, the mechanism behind this efficiency is still contested. A related question within the perceptual expertise literature is whether experts are storing semantic representations (or, the meaning) of domain-specific knowledge in memory, or perceptually specific representations (or, specific visual configurations and contours) of domain-specific knowledge in memory. As summarized by Reingold (2002), the current literature suggests that experts are likely relying on both of these mechanisms.

There is robust evidence to suggest that experts’ memory representations are perceptually specific (face perception: Richler & Gauthier, 2014; Gauthier et al., 2003; radiology: Nodine & Krupinski,
Krupinski et al., 1998; Waite et al., 2020; chess: Reingold et al., 2001). The question of whether this finding extends to music reading expertise is important for understanding the boundary conditions of expertise effects. Thus, in Experiment 2, I investigate the degree to which disrupting the perceptual specificity of a stimulus impacts music experts’ change detection.

As discussed earlier, the primary methodological motivator for the paradigm I used in my dissertation was used by Reingold et al. (2001). They observed change detection advantages for experts, only for legal game positions, and not for randomized positions (Reingold et al., 2001). This supports the conclusion that experts in chess possess perceptually specific encoding advantages relative to non-experts. Here, the extent to which chess stimuli were perceptually similar to legal configurations or rules that the expert had encountered prior impacted the degree to which encoding advantages by the experts were observed.

In Experiment 2, I implemented a rotation manipulation. Half of the stimuli were presented in the orientation that is familiar to musicians (i.e., an upright position; see Figure 5, Panel A, and half were presented in an orientation that is unfamiliar to musicians (i.e., a 90° rotation position) (see Figure 5, Panel B). A 90° rotation was selected because it optimally disrupted the configuration of familiar musical patterns within the stimulus. Unlike face perception studies, which often implement a 180° rotation to optimally disrupt familiar groupings of facial features (for a review, see Bilalić, 2017), rotating music scores by 180° fails to disrupt the contours and patterns of written music that are familiar to experts. Figure 5, Panel C illustrates the implementation of the flicker paradigm in Experiment 2. This flicker

9 Anecdotally, this 90° rotation proved to be frustrating for the expert participants, several of whom verbally expressed their frustration about the negative effect of this manipulation on their performance ability. At least two participants reported the strong desire to tilt their head in the headrest, in an attempt to compensate for the unfamiliar viewing orientation on rotated trials.
implementation is identical to that used in Experiment 1, except that I used a rotation manipulation rather than a complexity manipulation.

In addition to the change detection task, I implemented an automated operation span (or, OSPAN) task in Experiment 2 (Unsworth et al., 2005). The OSPAN task provides a measure of visual working memory capacity (VWMC; Conway et al., 2005). This task helps assess differences in VWMC between groups. This is important because efficient performance on the change detection task inherently involves VWMC. In order to complete the change detection task, experts retain in memory a percept of the original image, which they mentally compare to the changed image, in real time. If experts have a better VWMC than non-musicians, superior performance could be a result of this exclusively. Because visual change detection imparts a task demand that involves visual working memory (VWM), and because among experts, chunking alleviates VWM load (Thalmann et al., 2019), assessing differences in VWMC between groups was important to draw stronger conclusions about what factors support expert performance on a domain-specific change detection task.

**Experiment 2 Predictions**

Experiment 2 was designed to test the predictions of chunking and template theories that the performance advantage of experts will be specific to their domain of expertise.

To test the prediction of chunking and template theories that music reading expertise is perceptually specific, I implemented an orientation manipulation. As described earlier, evidence from radiology and chess suggests that experts’ superior performance on a domain-related task is not a result of overall superior perceptual abilities, relative to non-musicians. Experts do not possess superior perceptual abilities outside of their domain of expertise, relative to non-experts (e.g., Nodine & Krupinski, 1998 (radiology); Reingold et al., 2001 (chess)). An orientation manipulation is a straightforward and effective way to prevent experts from being able to rely on perceptually specific
memory representations to promote superior performance. To the extent that experts’ memory representations are perceptually specific, expertise effects should be eliminated or attenuated for trials containing music scores in unfamiliar orientations.

I tested the predictions of Experiment 2 with eye tracking measures (saccade amplitude and fixation duration), and accuracy and reaction time (RT) measures. I predicted that I would replicate my findings from the complex condition in Experiment 1, within the upright condition in Experiment 2. I predicted that I would observe robust expertise effects manifesting as faster RTs, higher accuracy, and fewer fixations, for experts relative to non-musicians. In contrast, I predicted an attenuation of expertise effects within the rotated condition, to the extent that experts’ memory representations are perceptually specific. I predicted that this attenuation would manifest in experts as longer RTs, longer fixation durations, and more fixations in the rotated (but not the upright) condition, reflecting more effortful processing. Specifically, I predicted that expertise effects would be minimized or eliminated in the rotated relative to the upright condition.

In order to assess whether expertise effects could be attributed to experts’ better memory overall, I used an OSPAN test to obtain estimates of participant’s visual working memory capacity (VWMC). Evidence of a better VWMC for experts would be observed if experts’ OSPAN scores were significantly higher than non-musicians’ OSPAN scores. Conversely, if the groups do not differ in their OSPAN scores, any differences between groups on eye movement/behavioral measures cannot be attributed to superior VWM among experts. I predicted that experts and non-musicians would not differ in their OSPAN scores, following the logic outlined above and the prediction that expertise is a result of deliberate practice, not innate talent (Ericsson et al., 1993).
Method

Participants

A total of 60 participants were recruited for Experiment 2. The expert participants were recruited from the University at Albany, SUNY, and the surrounding community, and the non-musician participants were recruited from the University at Albany, SUNY’s research participant pool (n = 30; 13 females). As compensation, experts received $10/hour and non-musicians received course credit. The experts had a minimum of 10 years of music reading experience (M = 12.7 years; range = 10–26 years), including music theory training and music performance of at least one instrument. The non-musicians were not able to read music. The experts’ mean age was 23 (median = 21; range = 18–42) and the non-musicians’ mean age was 19 (median = 18; range = 18–21).

The experts in Experiment 2 played a wide range of instruments: piano (12), guitar (7), flute (5), violin (5), trumpet (5), clarinet (4), bass guitar (4), trombone (3), alto saxophone (3), ukulele (2), French horn (2), cello (2), accordion (1), organ (1), soprano saxophone (1), tenor saxophone (1), viola (1), and ocarina (1). 90% of experts played more than one instrument.

All experts were actively engaged in playing a melodic instrument and reading sheet music at the time of the experiment, and nearly all experts (93%) had no lapses in their music training (i.e., had not stopped practicing and reading music since they began). All experts had studied music formally in an educational setting and reported having at least “advanced” knowledge of music theory. Most experts (83%) also studied music informally, by watching videos, completing composition and performance requests for others, or practicing “by ear” (i.e., listening to a piece of music and attempting to replicate it without the aid of sheet music). See Appendix A for the complete demographic questionnaire.
Materials and design

Building on the results of Experiment 1 where the largest and most compelling expertise effects were observed with complex (and not simple) music scores, only complex scores were selected as stimuli in Experiment 2. These experimental materials were generated from music scores composed during the Classical era and were drawn from a variety of instruments. Each stimulus contained four two-bar lines of treble clef music.\(^\text{10}\) All stimuli were roughly rectangular in size, with slight natural variations in absolute size.\(^\text{11}\) 104 unique stimuli were generated (100 experimental, four practice).

Half of the stimuli were presented in the orientation that is familiar to musicians (i.e., an upright position; see Figure 5, Panel A, and half were presented in an orientation that is unfamiliar to musicians (i.e., a 90° rotation position) (see Figure 5, Panel B).

As in Experiment 1, altered versions of each music score were also created to implement the flicker paradigm (Rensink et al., 1997). The altered version was identical to the original music score except that one music note or one chord (a chord is a vertical stack of music notes) was removed from the image. I manipulated orientation within-subjects, and used a 2 (Orientation: Upright or Rotated) × 2 (Expertise: Experts or Non-musicians) design.

Apparatus

Apparatus is identical to that of Experiment 1.

---

\(^{10}\) Stimuli in Experiment 2 were restricted to two bars per line (instead of four or more in Experiment 1) in order to maintain consistency across images, and to keep the image within the optimal range for the eye tracker (i.e., four-bar lines of music would be outside of the vertical trackable range in the rotated condition).

\(^{11}\) Upright orientation: height range = 7.7° to 9.6°; width range = 7.9° to 10.3°
Rotated orientation: height range: 7.9° to 10.3°; width range = 7.7° to 9.6°
**Procedure**

**OSPA**. The OSPAN task was administered after the participant signed a consent form and provided demographic information. Each trial included a math equation followed by a word (see Figure 6 for an example). The task was both to determine whether the math equation was correct by using the “Y” and “N” keys on the keyboard, and to read the word following the equation. After a set of math equation-word pairs, participants were instructed to recall all the words they saw in that set, in order by typing the words and separating them with a space. For words that could not be recalled, participants were instructed to type a “?” in that word’s ordinal place within the overall list. Set sizes of the math-equation word pairs randomly ranged from two to six items. Participants completed 66 trials, after which they were encouraged to take a short break before beginning the eye-tracking portion of the experiment.

**Change Detection Task.** Participants initiated each trial by fixating a gaze-triggered black fixation cross in the center of the computer screen. Participants completed a flicker paradigm task on each trial (Rensink, et al., 1997). Original and altered versions of the music score were presented briefly (for 700 ms each) and continuously alternated with a blank (i.e., white) screen presented between each alternation (for 100 ms). This presentation sequence continued either until a response was provided or two minutes had elapsed. To respond, participants fixated the location of the change while simultaneously pressing a response button on the gamepad. Once the button was pressed, their fixation location was recorded, and the trial ended.

Following calibration, participants completed four practice trials (two upright, two rotated) and 100 experimental trials (50 upright, 50 rotated). Presentation of stimuli in both orientations was randomly intermixed. Participants were encouraged to take breaks after each block of 25 trials. An important methodological advantage of Experiment 2 over Experiment 1 is that each music score was
presented in both conditions (counterbalanced across participants), but each participant only saw one version of each image. This allowed for maximal control, between conditions, over the slight naturally-occurring variations in the size of elements of the music score such as notes and expressive markings.

Results

I will first present the OSPAN results. Next, to examine the impact of expertise and orientation manipulation on change detection, I will report accuracy and reaction times (RTs), followed by three eye tracking measures (fixation duration, number of fixations, and saccade amplitude). For all measures reported below, I conducted 2 × 2 ANOVAs, with Orientation (upright, rotated) as a within-subjects variable and Expertise (experts, non-musicians) as a between-subjects variable. Error trials (i.e., trials in which the changing note was not accurately located) were removed from all analyses and constituted less than 1% of all trials. Experiment 2 data are displayed in Table 2.

**OSPA**

Table 3 displays the mean accuracy and t-tests for the OSPAN measure (Table 3, Panel A), as well as bivariate correlations of the variables “expertise”, “OSPA”, and “accuracy” (Table 3, Panel B). In support of my prediction, OSPAN scores were not significantly different across levels of expertise, \( r(58) = -0.11, p = .394 \). Relatedly, OSPAN scores were not significantly correlated with accuracy on the change detection task, \( r(58) = 0.09, p = .453 \). These results suggest that expertise effects observed in Experiment 2 are not a result of experts’ superior VWMC, as compared to non-musicians. Finally, expertise and accuracy were positively correlated, \( (r(58) = 0.30, p < .001 \). This reflects the finding that experts were more accurate than non-musicians overall in Experiment 2 \( (t(58) = -2.41, p = .019) \).
**RT and Accuracy**

Figure 7 displays the means and standard errors for the accuracy (Panel A) and RT (Panel B) measures. Experts were faster than non-musicians ($F(1,58) = 5.82, p = .019$). The upright trials elicited faster and more accurate responses than the rotated trials (RT: $F(1,58) = 83.50, p < .001$; Accuracy: $F(1,58) = 5.73, p < .001$).

In support of my predictions, expertise effects on performance were attenuated for rotated, but not for upright, trials. This conclusion is supported by a significant interaction of orientation and expertise for RT ($F(1,58) = 10.27, p = .002$). Follow up t-tests revealed significant expertise effects on RT in the upright ($t(58) = 3.03, p = .004$), but not the rotated ($t(58) < 1, p = .980$) condition. These t-tests also revealed significant expertise effects on accuracy in both the upright ($t(58) = -2.15, p = .036$), and the rotated condition ($t(58) = -2.31, p = .024$). Reaction times between groups were nearly identical in the rotated condition ($M_{\text{expert}} = 16734 \text{ ms}; M_{\text{non-musician}} = 16762 \text{ ms}$), reflecting the impact of the orientation manipulation on experts’ ability to complete the change detection task. This result is especially striking when observing the typical expertise effect on RT in the upright condition ($M_{\text{expert}} = 10816 \text{ ms}; M_{\text{non-musician}} = 13917 \text{ ms}$).

**Fixation duration and Number of fixations**

Figure 7 displays the means and standard errors for the fixation duration (Figure 7, Panel C) and the number of fixations (Figure 7, Panel D) measures. Relative to the non-musicians, the experts had shorter fixation durations ($F(1,58) = 17.05, p < .001$). There was no overall difference between groups for the fixation count measure ($F(1,58) < 1, p = .637$). Relative to upright trials, rotated trials elicited longer fixation durations ($F(1,58) = 94.16, p < .001$) and more fixations ($F(1,58) = 39.45, p < .001$).

Critically, expertise effects were eliminated in the rotated relative to the upright condition for the fixation count measure ($F(1,58) = 11.45, p = .001$). As there were no overall differences between
groups in either condition (upright: \( t(58) = 0.97, p = .336 \); rotated: \( t(58) = -1.72, p = .090 \)), the significant interaction is supported by the difference in fixation count for experts between conditions, \( t(29) = -6.98, p < .001 \). In support of the predictions of chunking and template theories, experts made significantly more fixations in the rotated condition (\( M = 39 \)) than in the upright condition (\( M = 28 \)). I interpret this result as evidence that experts were unable to rely on their perceptually specific memory representations to facilitate performance in the rotated condition, where experts viewed music scores in an unfamiliar and un-practiced orientation. Interestingly, non-musicians showed a small difference in fixation count between conditions (\( M_{\text{rotated}} = 34; M_{\text{upright}} = 31, t(29) = -2.01, p = .054 \)). I did not expect non-musicians to differ in terms of fixation count between conditions, so further research is needed to understand the implications of this finding. Unlike experts, non-musicians do not possess the requisite memory structures to interpret the written music in a meaningful way beyond basic pattern recognition. As such, changes in stimulus orientation are not as detrimental to non-musicians’ task performance as they are to experts’ performance.

**Saccade amplitude**

Figure 7 displays the mean and standard error for the saccade amplitude measure (Figure 7, Panel E). Saccade amplitudes were significantly larger for the rotated, as compared to the upright condition (\( F(1,58) = 74.53, p < .001 \)). As in Experiment 1, I observed no significant interaction between orientation and expertise for saccade amplitude (\( F(1,58) = 1.81, p = .185 \)). However, the saccade amplitudes of experts and non-musicians were more numerically similar to one another in the rotated condition (\( M_{\text{expert}} = 1.39^\circ; M_{\text{non-musician}} = 1.35^\circ \)) than they were in the upright condition (\( M_{\text{expert}} = 1.56^\circ; M_{\text{non-musician}} = 1.48^\circ \)). While these differences are not significant, they may be suggestive of increased effort by experts in the rotated condition. Future research could vary the size of the stimuli to explore if larger images (with greater numbers of music bars) would further magnify these numerical differences.
Experiment 2 Discussion

Experiment 2 tested the prediction of chunking and template theories that experts’ memory representations are perceptually specific (for a review, see Bilalić, 2017). In support of this prediction, I observed that expertise effects were much attenuated (and in some cases, eliminated) in the rotated, and not the upright, condition. Expert RT’s were nearly equivalent to those of non-musicians in the rotated condition, but RTs in the upright condition reflect an expertise advantage (i.e., experts are faster than non-musicians in the upright condition). Additionally, experts made significantly more fixations in the rotated condition, relative to the upright condition, $t(29) = -6.98, p < .001$.

As described earlier, evidence for the perceptual specificity of experts’ memory representations would be observed if experts did not demonstrate their typical performance advantages in an unfamiliar condition. The results of Experiment 2 provide support for the idea that music expert’s memory representations are perceptually specific. For experts, the unfamiliar condition prevents reliance on mental chunks to facilitate superior performance relative to non-musicians, and impedes experts’ ability to compensate for the increases in task demands introduced by the rotation manipulation.

The OSPAN results revealed no correlation between level of expertise and OSPAN score, $(r(58) = -.12, p = .394)$. This finding suggests that there is no significant difference between the visual working memory capacity of experts and non-musicians, providing compelling support for the conclusion that the expertise effects observed within the change detection task cannot be attributed to overall group differences in working memory capacity. Rather, it is more likely that experts’ superior performance is supported by domain-specific knowledge, and not a larger VWMC.

General Discussion

In the general discussion of my dissertation, I will first provide a summary of the main findings, followed by a discussion of the significance of this research, contextualizing my findings in terms of their
methodological, theoretical, and empirical importance. I will conclude this section of my thesis with a brief discussion of some additional questions that arise out of the current work, related to the myriad ways that “expertise” can be defined, how the physical characteristics of written music may promote “musical chunk” formation, and the effects of practice on change detection.

**Summary of Main Findings**

In Experiment 1, I manipulated visual complexity to test the prediction that experts’ domain-specific knowledge permits them to compensate for increases in visual complexity. In support of this prediction, non-musicians were more greatly affected by increases in visual complexity than were experts (e.g., smaller saccades and longer fixations durations for complex than for simple trials). In Experiment 2, I manipulated the familiarity of the music’s visual configuration by contrasting upright (i.e., typically oriented) scores with those that were rotated by 90° to test the prediction that visual expertise is perceptually specific. In support of this prediction, I observed no expertise effects on eye movements in the rotated condition, while observing robust expertise effects in the upright condition, as in Experiment 1. Taken together, these findings support the conclusions that music reading experts may engage in more parafoveal processing than non-musicians, and display perceptual advantages that are remarkably perceptually specific. Consistent with the assumptions of chunking and template theories, these results suggest that expert musicians can use chunking mechanisms (e.g., perceptual grouping) to facilitate encoding during domain-specific change detection.

Consistent with these predictions of chunking and template theories, and similar findings in other domains like chess and medicine, my dissertation results suggest that expert musicians can use chunking mechanisms (e.g., perceptual grouping) to facilitate encoding during domain-specific change detection. These findings also align with recent others in the domain of music reading supporting the
conclusion that music reading experts are relying on mental chunks to rapidly identify relevant
information (i.e., Maturi & Sheridan, 2020).

**Significance of The Work**

My dissertation work contributes to the perceptual expertise literature in multiple ways:

**Methodological Contributions.** My dissertation work contributes a novel paradigm to the
growing number of methods that researchers can use to investigate music reading expertise. This
paradigm could be even further adapted for use in different unique domains, such as lifeguarding
expertise (e.g., Page et al., 2011; Langendorfer & Beale-Tawfeeq; 2022) or handwriting expertise (e.g.,
Fears & Lockman, 2018). This application could shed light on the bigger picture of expertise effects on
change detection. Moreover, this application could also help clarify the conditions under which, and the
domains of expertise wherein, chunking facilitates domain-specific task performance.

Additionally, the change detection task that I used in my dissertation research did not require
prior music knowledge in order to complete with a high degree of accuracy, because the music was
isolated to its visual component. This allowed for equal comparisons between groups. The accessibility
of this task to true non-musicians who could not read music is a strong methodological advantage of this
work. Additionally, investigations in the domain of music reading expertise permitted me to obtain large
and equal sample sizes of experts and non-musicians. This methodological strength is unique to the
domain of music reading; other expertise studies sometimes have comparatively small sample sizes of
experts (i.e., \( N < 10 \)).

**Theoretical Contributions.** My dissertation work contributes theoretically to larger
conversations within the perceptual expertise literature. As discussed earlier, current research efforts in
the field are focused on developing a more integrated theory of perceptual expertise (for a related
discussion, see e.g., Brams et al., 2019). This necessitates the inclusion of novel domains of expertise
into the conversation, in order to comprehensively understand the boundary conditions of the range of perceptual expertise effects that are observed throughout this literature. My dissertation work provides support for the idea that the domain of music reading expertise is well-suited for exploring perceptual expertise effects with a change detection paradigm.

**Empirical Contributions.** My dissertation work constitutes a meaningful empirical contribution to the field of perceptual expertise research. Across two experiments, I observed conditions under which music reading expertise effects are magnified (i.e., visual complexity in Experiment 1), as well as conditions under which these effects are minimized, and in some cases, fully eliminated (i.e., orientation manipulation in Experiment 2). These comparisons are important because they represent strong empirical support for the idea that chunking mechanisms underlie the performance advantages of music reading experts.

**Defining Expertise**

In my thesis, I defined an “expert musician” as someone who had at least ten years of music reading experience. I arrived at this definition after conducting a review of relevant literature, but there is surprisingly little agreement on the factors that underlie “expert musicianship”. This stands in stark contrast to the domain of chess, which makes use of the algorithmic Elo rating system. The lack of agreement on definition is likely due at least in part to the multimodality of music; while I defined an expert musician based on their ability to read music, it is important to note that reading sheet music is not a necessary condition for music performance at an expert level. There are many 20th- and 21st-century music icons who have achieved international success without the ability to read music, such as Jimi Hendrix, Bob Dylan, and Taylor Swift; questioning these musicians’ expertise on the basis of music reading ability seems like a weak argument in the face of their successes.
A contemporary meta-analysis of the music psychology literature observed three primary ways that researchers in this field define a “musician” (note that this may be subtly different from the definition of “expert”) (Zhang et al., 2020). One component of musicianship is “skill”, which is often operationalized as “years of music training”. This is most similar to my definition in this current work. Another component of musicianship is “identity”, which can be measured via self-report to demographic questions like “I consider myself a musician”. The third way that musicianship is defined within the literature is through “predisposition”, which can be measured with self-report (“I can identify a given note”) or some variant of a musical aptitude test (e.g., having participants judge or name musical chords) (Zhang et al., 2020). Across 95 papers that were included in the meta-analysis, 51% defined musicianship by years of training, and the general consensus was that a “musician” had at least six years of training. This estimate is lower than the 10 year definition in my dissertation, which may reflect some of the lack of overlap of definitions between separate fields studying similar constructs (i.e., “expert” vs. “musician”).

Amidst the multitude of ways that expert musicianship can be defined, I believe that studies using an eye tracking methodology have great potential to reveal key factors underlying this skill. As mentioned earlier, eye movements are required to align the fovea with a specific region of a visual scene or image; this highlights the link between eye movements and attention. For this reason, during complex visual-cognitive tasks, like music-related change detection, a strong link between eye movements and cognition has been documented (i.e., the eye-mind link; Reichle & Reingold, 2013). As such, eye tracking provides a useful index of ongoing cognitive and perceptual processing by providing fine-grained information about how attention is allocated as a function of time (for a related discussion, see Sheridan et al., 2020). Eye tracking methodology also has the benefit of being a covert measure, and is not as impacted by potential confounds or bias involved with other dependent measures, such as self-report or a basic behavioral task requiring a button press for response (i.e., there is inherently a
millisecond-length delay between the identification of a target stimulus, and subsequent response to the stimulus). Because eye tracking has the ability to reveal information that the participant cannot verbalize, or is not aware of, it holds great promise for exploring the specific components of behavior that promote or reveal expertise effects (for a related discussion, see Reingold & Sheridan, 2011).

A more holistic (and therefore, accurate) approach to defining music expertise must be sensitive to the multimodal nature of music. My dissertation research isolated the skill of music reading to its visual component; this is a strong methodological choice, as I avoided aural confounds present in many other music reading studies (i.e., pieces with faster tempi are necessarily more complex), but it is also reductionistic. I arrived at my decision to use a ten-year minimum of music reading ability for inclusion in the expert group in both experiments by conducting a literature review of relevant eye-tracking studies, most of which are cited within this thesis. Inclusion criteria varied somewhat drastically across previous experiments in the literature, but at least eight years of experience was typically the consensus. I decided to use a more conservative criterion of ten years to maximize the likelihood of observing true differences between groups, should they exist. Lastly, many of the demographic measures were highly correlated among experts (i.e., most experts had extensive formal and informal training, played multiple instruments, etc.). Given this, it seems reasonable to question how much of the variance these demographic measures independently account for. In the spirit of capturing the unique multimodal nature of music expertise, future studies should inquire about the amount of time spent both performing music and improvising music, in addition to more specific questions about the type/quality of practice, perhaps even including a verbal self-report of a successful practice session. Ultimately, I think that the definition of “music expert” must both be relevant to the nature of the task at hand/modality being studied, and be sensitive to the complex multimodal nature of music expertise.
**Musical Chunks**

While my thesis provides support for the idea that expert musicians are engaging in chunking during domain-specific tasks, this begs the question of what exactly constitutes a musical chunk? To what specific physical elements of a given stimulus are experts attending when completing a music-related change detection task? I discuss these issues with evidence from other studies in the Future Research Directions section of this thesis, but additional explorations into my current data have the potential to provide significant insight into these questions.

In any perceptual task, the ease or difficulty of locating a target within a given trial is impacted by a multitude of factors. Specifically, in my studies, physical aspects of the target note/note stack, the relative location of the measure containing the target note(s), and individual familiarity with specific target note(s) likely influence the salience of that target and its subsequent detection (or detection failure). “Salience” is also heavily related to what participants are actually doing during a given change-detection trial; additional eye movement analyses can help shed light on these questions (i.e., scan path analyses discussed in Future Research Directions). As mentioned in the immediately preceding section, eye tracking methodology holds great promise for disentangling these separate factors impacting expert change detection.

With the help of a research assistant, I categorized all stimuli from both experiments on 18 different dimensions related to the physical characteristics of the stimulus, such as whether the target note has an expressive or accidental marking, whether the target note occurs on a downbeat, and whether the target appears to be part of a highly familiar “chunk”, like an arpeggio. While the exploratory nature of this kind of analysis makes it challenging to draw strong conclusions for the current data, understanding what elements of a stimulus promote or draw expert attention could be acutely beneficial to the intelligent design of future experiments investigating perceptual expertise.
This coding helps address a number of predictions about the specifics of expert eye movement behavior during a change-detection task. For example, any changes that are anomalous or unexpected should be harder for an expert to detect, much in the same way that spelling errors are harder to detect in high frequency words (Rayner & Raney, 1996), and that shorter high frequency words are sometimes overlooked altogether (Healey, 1976). Additionally, expectations about familiar melodic runs or arpeggios should influence change detection, such that a note missing from the middle a familiar sequence of notes, rather than the beginning or the end, should be more easily detected (because expert musicians have expectations about the locations of specific notes in familiar phrases). I would also predict a difference between the ease of single note vs. chord change detection, that is correlated with the type of instrument an expert plays. For example, those who play polyphonic instruments (i.e., capable of playing multiple notes simultaneously, like the violin or piano) may be less sensitive to changes occurring within chords than changes occurring for single notes (again, related to expectancy and familiarity).

A more exclusive exploration of these stimulus characteristics is one of my primary future research directions for my dissertation work, with the long-term goal of defining the boundary conditions of what constitutes a “musical chunk”.

**The Effects of Practice on Change Detection**

One final point of general discussion involves the question of practice effects, or the observation that participants’ performance improves over time during an experiment because of practice with that specific task, and not because of skill level or some inherent difference between groups (Donovan & Radosevich, 1999). In order to assess whether there were practice effects occurring within my experiments, I isolated the last 1/3 block of trials and conducted identical analyses on these data, in order to compare this to the ANOVA for all trials.
The analysis of the last block of trials yielded results that were very similar to the analysis for all trials, as the numerical patterns are the same for all interactions in both experiments. Broadly speaking, I observed a main effect of faster processing in the last block (i.e., slight reductions in RT for both groups in both experiments), but I also observed a reduction in accuracy over time. These results likely reflect a speed-accuracy tradeoff, where participants become faster over time, but compromise accuracy in the process. However, faster processing may reflect a “true” practice effect. These interpretations are made with caution; given the exploratory nature of this post-hoc analysis, the fact that the current work was not designed to detect practice effects, and because statistical power is reduced when restricting a range, more formal investigation is needed to determine the presence or absence of practice effects during music-related change detection.

In the next section of my thesis, I will discuss some limitations to the current work as well as potential future research directions for my line of dissertation research.

**Limitations and Future Research Directions**

**Limitations**

Although my dissertation focuses on the theoretical framework of chunking and template theories, this is not the only perspective in the visual expertise literature. For example, it is unclear whether support for **holistic processing** is also observed within the current data. Sala & Gobet (2017) reported compelling evidence in support of chunking (and specifically, template) mechanisms, in lieu of holistic processing mechanisms, as an explanation for perceptual expertise effects. While it is true that expertise effects for random (relative to familiar/legal) visual configurations are much attenuated (e.g., Reingold et al., 2001), chess experts do retain a small (but reliable) performance advantage relative to non-experts when viewing randomly configured stimuli. In a meta-analysis spanning multiple domains of expertise, Sala and Gobet (2017) concluded that even within random visual configurations, some
meaningful patterns do still occur by chance, which experts are more likely to attend to than non-experts (for a related discussion, see Gobet & Simon, 1996). Critically, this result can be easily accounted for by chunking and template theories, but not by holistic processing theories, because random visual configurations do not contain overarching structures, or “wholes” (Sala & Gobet, 2017). This paper in particular provides compelling evidence in support of chunking and template theories as a theoretical foundation for multiple domains of expertise.

Other research suggests that musical phrases can be processed in a holistic way by both experts and non-musicians but concludes that this holistic processing occurs for different reasons between groups (Wong & Gauthier, 2010). Using a sequential matching paradigm, participants decided whether two sequences of notes were identical or different, while the position of a target note within a note sequence was manipulated. Wong and Gauthier (2010) reported that expert holistic processing reflected automatic allocation of attention to all notes in a sequence, while non-expert holistic processing reflected different strategies that fulfilled the task demand, given their inability to process all of the notes at the same time. Future work should aim to dissociate chunking and template theories from holistic processing frameworks, as it is clear that both mechanisms can contribute to an understanding of observed expertise effects among music readers. Studying the modulation of holistic processing effects by different factors may therefore be important for revealing the mechanisms that support these expertise effects (for a related discussion, see Richler et al., 2011).

Another limitation of my current work is that it is restricted to visual expertise, which is but one aspect of the music expert’s skill set. Music expertise involves input from multiple senses, and it is challenging to address all of these factors within one experiment or set of experiments. My dissertation work represents an exploration of the generalizability of chunking and template theories to a novel domain, with an emphasis on the visual component of written music. While this is a methodological
strength of my work, as potential confounds related to the auditory domain (such as *tempo*) are eliminated, it also necessarily restricts the scope of my conclusions.

**Future Research Directions**

One avenue for future research involves further analyses of the existing data. Like expert chess players, expert musicians may sometimes fixate on blank spaces in a music score (or a chess board, for chess experts). Specifically, two studies that have analyzed fixation location noted that music experts were almost equally likely to fixate on a group of notes as they were a blank space in the music score (Truitt et al., 1997; Gilman & Underwood, 2003). Similar findings have been observed in the domain of chess (e.g., Reingold & Sheridan, 2011). These findings support the conclusion that experts have a larger visual span than non-experts or are able to perceive more visual information in a single fixation than are non-experts. By fixating a blank space on a chess board or within a music score, experts may be able to perceive more and larger visual patterns in their parafoveal or peripheral vision, thereby supporting an expert advantage. Conducting analyses on the scan path of an experts’ eye movements may help explain the reason for fixations on blank space, exploring the variable of “fixation location over time” (Tatler et al., 2005). Finally, analyses could be conducted on the eye movement data of participants who have more music reading experience than a non-musician, but less than an expert. Analyzing these data would permit the treatment of expertise as a continuous, rather than a dichotomous, variable.

Relatedly, in addition to the expert and non-musician data presented in this thesis, I also collected substantial intermediate data (at least 15 intermediates per experiment). These individuals had more than 1 year of music reading experience, but less than 10 years. Dichotomizing the variable of “expertise” is methodologically useful, but necessarily eliminates variability within, and richness of the potential differences between, groups. Including eye movement data from participants of varying skill levels, and treating expertise as a continuous variable, could provide significant insight into appropriate ways of defining “music expertise”.

52
Another future research direction involves the direct application of the flicker paradigm used here to other types of music expertise. As mentioned earlier, not all music experts can read sheet music. As such, this dissertation’s flicker paradigm could be extended for use with different genres of music (not just Classical), as well as with different visual components involved in music expertise beyond reading sheet music (i.e., instrument identification and classification). Along this vein, more detailed questions about interaction of instrument type with change detection ability could be asked. Future experiments could modify the instrument type of the stimuli used (here, single-line pieces were used exclusively, eliminating pieces with multiple lines of music, e.g., a piano with a treble line and a bass line). The impact on task performance of the instrument most familiar to expert participants could also be explored. To the degree that an expert is familiar with the pieces used as stimuli (if they are not created uniquely for the purposes of the experiment), I would predict that change detection within a flicker paradigm should be facilitated. Moreover, nuanced questionnaires about the specific types of instruments played could provide enough information to conduct a regression analysis exploring the amount of variance that could be explained by expert’s variable knowledge of instrument type.

This flicker paradigm could also be applied within other domains of expertise, such as radiology. Attempts at replication across domains using identical methodology would directly address questions of generalizability of these perceptual expertise effects. This would be particularly interesting within the field of radiology, because expertise effects in this field are generally best supported by holistic processing theoretical frameworks (Sheridan & Reingold, 2017), and not by the chunking and template theoretical framework.

It would also be interesting to explore the difference between legality and familiarity as it relates to written music. While no musical performance (i.e., auditory) can be “illegal” per se, written music (i.e., visual) can contain illegal locations or omissions of notes (e.g., 5 whole notes written in a measure marked with a 4/4 time signature) that would violate the rules of Western music theory. It
would be interesting to manipulate both musical legality and musical familiarity within the same experiment using eye-tracking methodology and an expertise manipulation. Future experiments in other domains could consider the idea that these two concepts may not be synonymous, and could help disentangle the separate effects of both factors on eye movement behavior.

One additional avenue of future research would be to explore replication of music reading expertise studies that are not (at present) supplemented by eye movement data. Because eye movement measures are especially useful for testing predictions of the chunking and template theoretical framework, collecting this additional data would be especially helpful for comparison. Two other studies that I discuss below suggest exciting avenues for future research related to the nature of experts’ mental representation of stimuli within their domain, and what constitutes a written music “chunk”. While their results appear to be consistent with chunking and template theories, it is not possible to fully assess support for this prediction without supplemental eye movement data, which offers an exciting and approachable avenue of future exploration.

In one study, Brodsky and Kessler (2017) implemented a music-related variant of the Stroop paradigm where both expert and non-musicians attended to one of two features of a simple musical stimulus (namely, the perceived direction of the notes and the beam connecting them) and made a congruency judgment on a given trial. Importantly, their first experiment was unimodal (i.e., visual input only), and the second was bimodal (i.e., visual and auditory input). During Brodsky and Kessler’s (2017) first experiment, experts showed a larger Stroop effect relative to non-musicians when asked to attend to the beams of the musical stimulus, but not the notes; the exact opposite pattern of results was observed for non-musicians. While the authors suggest that this result makes sense in light of the fact that experts “spend their lives decoding pitch heights” (p. 198) (i.e., and not beam slopes), results from their second experiment provide insight into this interpretation.
As a basis of their domain-specific knowledge, many music experts rely on the conventional rules of Western music theory that dictate the shape (or contour) of a written melody by the nature of its component notes (Clendinning & Marvin, 2016). For example, when writing eighth-notes higher than an A4 on a treble clef, conventional rules dictate that the beams connecting these eighth-notes be written below the notes, rather than above. During their second experiment, Brodsky and Kessler (2017) reported that experts’ ability to perform the musical Stroop task significantly declined when one of the visual features of the stimulus (i.e., either the notes or the beam) were incongruent with the auditory feature of the musical stimulus. Importantly, this pattern of data was different for non-musicians, who showed no differences in overall RT across all conditions, regardless of visual/auditory congruency.

This study may reflect the perceptual specificity of music reading expertise; it may be the case that experts’ extensive practice with legal (i.e., congruent) musical configurations contributed to their greater impairment relative to non-musicians, especially with the inclusion of auditory input (that could itself be either congruent or not). While it is likely true that music reading does not often involve the direct interpretation of beam slope, this variable still had a significant impact on experts’ task performance, attenuating their RT to approximate non-experts’ while also suggesting that beams between notes provide important information for the expert that extends beyond the basic grouping of musical notes (Brodsky & Kessler, 2017). A review echoes this point about the utility of musical beams, suggesting that “…perhaps beamed examples are chunked perceptually and thus can be processed in fewer fixations than nonbeamed examples” (Madell & Hébert, 2008, p. 164-5). Applying an eye tracking methodology in an experiment with a similar design may provide information about the role of musical beams in the formation of musical chunks.

Broadly speaking, Brodsky and Kessler (2017) capitalize on the multimodality of music in their experiment – a unique aspect of the domain of music that merits further exploration. Experts’ auditory imagery can interfere with the recognition of studied melodies, especially when hearing a melody that is
incongruous with a melody being read (Brodsky et al., 2003). Some researchers have suggested that auditory imagery disrupts a form of phonological encoding among experts, and is related to subvocalization of the melodic phrases themselves (Schürmann et al., 2002; Smith et al., 1992; Smith et al., 1995). Future research exploring the interactions between the visual and auditory modality within the domain of music reading expertise will produce a more holistic picture of the music expert’s skill set than would exploring either modality in isolation.

The second study that could benefit from a replication attempt with supplemental eye movement data is Halpern and Bower (1982). These authors provide compelling evidence in support of the perceptual specificity of music expertise. Inspired by the work of Chase and Simon (1973a), these authors created simple melodies that varied in terms of their musical familiarity (i.e., “good”, “bad”, and “random”; this was determined based on the results of a pilot experiment where experts judged the melodies on their compositional quality). A separate group of experts and non-experts viewed each 10-note melody for 4 seconds, and recalled the melody from memory (i.e., hand-wrote the melody onto a set of blank measures). Halpern and Bower’s (1982) results are quite similar to those of Chase and Simon (1973a), such that expert performances outperformed non-experts on “good” melodies, but their recall declined as melodies became less familiar. Critically, while the difference between expert and non-expert recall in the “random” condition was significant ($p < .01$), it was the smallest difference between groups across conditions. In other words, expert musicians’ recall of domain-specific configurations was impacted by the familiarity manipulation to a much greater extent than was non-experts’, supporting the idea that experts’ memory representations are perceptually specific. Patterns of eye movements that would support these conclusions would include increased fixations on blank space between notes for “good” melodies, and not “random”, as well as fewer fixations and longer saccades in conditions that were more familiar to the expert.
It may also be the case that spacing between musical chunks supports expert performance, especially given that expert musicians are quite sensitive to the spatial relationships within the music when reading it (e.g., the spacing between a note’s head and the measure/bar line; Aiba & Sakaguichi, 2018). For example, Stenberg and Cross (2019) asked whether the inclusion of “of white spaces as a chunking device [could] facilitate of the recognition and processing of musical symbols” (p. 2) – a music-related interpretation of interword spacing in written text, the inclusion of which is known to support perception in text reading (Bassetti & Lu, 2016; Hsu & Huang, 2000). Stenberg and Cross (2019) separated parts of the musical score with small white vertical gaps, and reported evidence consistent with this claim among expert music readers. Experts’ sight reading performance was better (i.e., higher accuracy) in conditions where the white gaps separated parts of the music that was neither too small (i.e., single notes) nor too big (i.e., entire musical phrases). Rather, task performance was best when the gaps separated an “optimally sized” chunk (i.e., three or four notes of a melody underneath a phrase marker). Given the results described here, it seems that internote spacing is yet another promising variable for further investigation in relation to music reading expertise.

Other suggestions for future research involve implementing different methodologies to study music reading expertise, such as the think-aloud paradigm (i.e., the process of verbally eliciting otherwise silent ongoing thought processes during a problem-solving task; e.g., Ericsson & Simon, 1984), the moving-window technique (i.e., restricting a participants’ gaze to a pre-defined window, such that the participant can only clearly see what is directly in their foveal vision; e.g., McConkie & Rayner, 1975), and co-registration (i.e., the utilization of eye tracking in conjunction with EEG methodology). The unique and complimentary contributions of the components of these methodologies are promising tools for researchers interested in parsing the multifaceted nature of music reading expertise. Co-registration may be a particularly promising avenue for future research, given the unique ways in which eye tracking and EEG methodologies complement one another (Baccino, 2011). Specifically, while eye tracking
provides fine-grained spatial information about behavior during a trial, EEG output (i.e., ERPs) can provide fine-grained temporal information (Hillyard & Kutas, 1983). During many previous EEG studies, free visual inspection was not possible, because the EEG record is contaminated by ocular artifacts like sequences of saccades. For this reason, co-registration has proved to be especially useful in natural reading studies (e.g., Dimigen et al., 2011), as the eye movement record can be compared against the EEG record to isolate and/or exclude ocular artifacts during analysis (Baccino, 2011).

Future work specifically using eye movement measures to understand the impact on performance of the features of written music has the potential to shed light on the relatively ubiquitous challenge of learning to read written musical notation (Haug, 1990; Saxon, 2009). Somewhat paradoxically, some of the literature reviewed here suggests that the ability to read musical notation is a critically important, but often overlooked, aspect of music performance and expertise (e.g., Penttinen et al., 2013). While in some cases, sight reading is not required of a music expert (for a related discussion, see Sloboda, 1984), multiple reviews suggest that this skill is “an integral part of the musical experience for all musicians” (Wristen, 2005, p. 1), and that anyone wanting to become a musical expert “will most likely have to master music notation” (Lehmann & Kopiez, 2009, p. 1). When simultaneously considering both the difficulty and importance of music reading to the aspiring music expert, the incentive and desire to improve this skill becomes apparent.

Some literature suggests the use of music notation “just as a performance aid” (Penttinen et al., 2013, p. 212-3) in a classroom setting inherently links the extraction of musical information to the slow learning of written music. In other words, it may be the case that the lack of instructional emphasis on the music notation itself contributes to a failure to acquire sight reading fluency (Penttinen et al., 2013). Some authors even argue that this struggle may prevent countless music students from pursuing the skill entirely, potentially significantly contributing to students dropping out of music lessons (e.g., Mills & McPherson, 2015). It is indeed the case that relatively few individuals who begin studying music reading
truly master the skill of sight reading (e.g., Gudmundsdottir, 2010; Zhukov et al., 2016), even at a moderately advanced level (e.g., Zhukov, 2014).

Importantly, some contemporary research suggests that there may be considerable costs associated with “mastering” the skill of reading written music notation. For example, Classically-trained musicians are significantly less able to play melodies by ear than musicians with other, less “formal” training backgrounds, such as improvisational jazz musicians (Corcoran & Spiro, 2021; Woody & Lehman, 2010). Neuropsychological work revealing less activation in the right auditory cortex during a music performance task for Classically-trained musicians relative to their improvising peers bolsters these behavioral findings (Harris & de Jong, 2015). These findings imply that music students’ “score-dependency” (i.e., an over-reliance on written music notation to the detriment of other important parts of musicality, such as playing by ear) reflects “an extreme form of overlearning” (Corcoran & Spiro, 2021, p. 73) that disadvantages the young musician.

This area of study also considers the importance of the student’s own experience and confidence level during music training. A student who lacks self-efficacy, which can certainly be caused by struggles associated with learning how to read written music, may be less inclined to pursue this skill development entirely. Research in this vein reports greater performance gains for students in an aural instruction group (i.e., learning to play an instrument without an explicit focus on notation) relative to those in the notation instruction group, as well as significantly greater post-treatment self-efficacy (Watson, 2010).

The work described here takes a much broader interpretation of “music literacy”, with consideration to the benefits that music students reap from “deep engagement with notated and un-notated music” (Abrahams, 2021, p. 81). Echoing this point in his book about the interaction of aural and visual music literacy, Edwin Gordon writes, “what is not seen in notation is far more important than
what notation can accommodate” (Gordon, 2004, p. 20). This quote emphasizes the fact that the richness of the experience of playing music (both written notation and improvisation) cannot be fully communicated solely via written ink on a page. The work described here may shed light onto potential reasons why benefits of chunking for learning to read written music notation are often not observed in studies focusing exclusively on the practice of written music (e.g., Pike & Carter, 2010).

To conclude my Future Research Directions section, I will briefly discuss some of the broadest implications of my work. At its core, my thesis research is concerned with the question of human perception. How are changes within complex and dynamic visual scenes detected? What (if any) stimulus features best promote rapid and accurate detection? How does expertise level interact with change detection? These questions are also relevant to applied fields of study, such as machine learning (e.g., Laird et al., 1986), cybersecurity (e.g., Xu et al., 2016), and artificial intelligence (AI). Many parallels exist between the tenets of chunking and template theories, as described within this thesis, and the principles of Music Information Retrieval (MIR) research (Downie, 2003; Schedl et al., 2014). The field of MIR is concerned with the “extraction and inference of meaningful features from music” (Schedl et al., 2014, p. 2), much in the same way that experts extract meaningful features from a music score using chunking techniques. MIR also involves indexing of these meaningful features for future use, e.g., music recommendation systems within widely-accessible music streaming services (e.g., Spotify) (Schedl et al., 2014). There may be similarities between this indexing, and the way that experts rely on mental chunks and templates to promote performance on new domain-specific tasks. These parallels between fields of study investigating similar core questions have the potential to foster more enriched and dynamic conversations about the nature of human perception, and enhance future research and scholarship.

At this point, I will turn my attention to a more speculative discussion of some broader connections that I see between my present dissertation work and my other interests as a college-level professor and cognitive psychologist.
Broader Connections

One of my primary interests as a researcher and an educator is understanding how principles of cognitive psychology can be leveraged in an applied way to manifest in meaningful changes for students in the college classroom. Recent popular-press books, such as “Make It Stick: The Science of Successful Learning” (Brown et al., 2014) emphasize using these principles to promote deep learning and enhance information retention. More narrative works, such as “What the Best College Teachers Do” (Bain, 2004), touch on these principles without naming them as such. As an example, the author suggests that “[t]he best teachers allow students to try, to fail, and try again” (p. 4); this can be viewed as an application of the testing effect (i.e., final test performance is better enhanced by repeated testing than by repeated re-study opportunities). Most re-study opportunities do not permit the student a chance to “fail”.

Speculating about the connections between my dissertation work and these professional interests has led me to consider questions such as the following: Can the principles of chunking and template theories be applied to enhance music reading students' learning? The answer to this question necessitates an understanding of what exactly constitutes a musical “chunk”. While contemporary reviews of the music expertise and eye tracking literature suggest that these data do “seem to reveal some evidence for chunking strategies” (Madell & Hébert, 2008, p. 165), our formal understanding of what constitutes a musical chunk appears to be lacking. However, a recent thesis from my lab (i.e., Maturi, 2021) examined this exact question with eye tracking methodology. In her dissertation, Maturi (2021) observed that expert musicians spent a longer proportion of time viewing relevant regions of a target stimulus, as compared to non-musicians. These longer and less frequent fixations can be interpreted as support for the idea that expert musicians can rapidly focus their attention on the most important or relevant parts of a domain-specific stimulus. Understanding the specific importance of these longer-fixated regions likely holds at least partial answers to the question of what comprises a musical chunk, from an expert’s perspective.
Other than this contemporary work, I located only two papers that formerly discuss the components of musical chunks, and ways that they could be leveraged to improve music reading ability– specifically, sight reading ability (i.e., Haug, 1990; Saxon, 2009). Sight reading is one of the many ways that a musician may engage with their domain and practice with the goal of acquiring expertise. These papers come from the music education sphere and make no reference to either Chase and Simon (1973a) or Gobet and Simon (1996). This observation represents a lack of cross-talk between the domains of perceptual expertise and music education.

Haug (1990) and Saxon (2009) make some recommendations about how to consider musical chunks when teaching how to read music. The authors advocate for an emphasis on teaching intervals rather than note names to discourage sight reading in a note-by-note fashion. They also suggest that more advanced students could engage in practice with mental chunks through the process of transposition (or, the on-line process of mentally altering the pitch height of the music as it is being played). Transposition activities have the potential to focus attention on visual patterns, in addition to highlighting the pitch patterns that comprise the contour of a melody. Reading musical chords from the bottom up is also recommended to highlight musical chunk structure; the bottom note of a musical chord may contain more information about the piece of music as a whole (e.g., key signature) than any of the other notes (Haug, 1990; Saxon, 2009). A recent dissertation even posited three different “types” of musical chunks: micro (i.e., intervals and rhythms), macro (i.e., larger phrase structures and themes within the music), and stylistic (i.e., accidentals) (Teo, 2020). Considering the myriad ways that written music can be perceptually grouped could help further our understanding of what constitutes a musical chunk, as well as our understanding of the different ways that music may be chunked by groups with varying levels of expertise.

Importantly, the chunking and template theoretical framework for interpreting perceptual expertise effects would suggest that students lack the cognitive or mental structures that experts
develop over hours of deliberate practice within their domain of expertise. Non-musicians’ lack of these mental templates renders any learned chunks inherently less useful for this demographic. This connection appears to be lacking within the music education work reviewed here, and I believe that this is important.

In terms of understanding how chunking behavior may benefit someone learning the skill of sight reading, a small body of literature has manipulated the presence of chunking-based interventions with the sole purpose of improving participants’ ability to visually perceive and sight read a music score. In these papers, “chunking interventions” were used where, after random assignment, participants were instructed how to engage in a specific chunking behavior multiple times in a classroom environment for the duration of the treatment period. Pre- and post-test scores established a baseline for evaluating sight reading performance at the time of final test (Gaynor, 1995; Hagen et al., 2012; Pike & Carter, 2010).

Gaynor (1995) used a manipulation designed to teach effective eye movement behavior as an aid to chunking. This author used a “chunking” technique intended to approximate the eye movement patterns of experts, such as instructing non-musicians to read alternating staves of written music and blank measures. The blank measures were presumably inserted in an effort to encourage longer saccades among the non-musicians (which has been established as an individual indicator of perceptual expertise), thereby encouraging fixation on the next measure containing written music before they had finished performing the prior measure. This measure appears similar to the EHS mentioned earlier.

Gaynor (1995) observed no overall benefit of chunking, but did report a significant benefit of chunking for non-musicians only, as evidenced by a significant increase in post-test (relative to pre-test) scores. This result may not be particularly surprising, especially if music experts are already engaging in
chunking behavior to aid their performance. While it is not possible to draw strong conclusions from these results, they reflect a clear need and desire within music education to improve sight reading skills.

Another study applied similar “eye guidance” principles in an effort to improve sight reading ability between groups after treatment (Hagen et al., 2012). Technological advances permitted more nuanced manipulations during a sight reading task, such as a “sweeping thin vertical bar that move[s] across a measure in time” (p. 234), and using color to highlight either an entire measure of music or individual notes (again, updating appropriately in time with the music’s tempo). At post-test, no differences were observed between groups, but importantly, significant differences within each group were observed, as evidenced by more accurate rhythmic and note execution during post-test than pre-test. At this time, it is not possible to know whether these results can be attributed to a particular experimental manipulation, or some other variable. However, this study is still interesting because it provides further suggestions for ways that chunking techniques may be implemented in an effort to improve sight reading ability.

Pike and Carter (2010) investigated whether an improvement in sight reading ability could be observed, using certain chunking techniques to emphasize certain rhythmic or pitch patterns (such as circling the musical notes that comprised the main melodic line) in a similar pre-post test design as Hagen et al. (2012). Again, no benefit of chunking was observed; no differences between treatment groups were observed, although all groups improved with practice. Pike and Carter (2010) explicitly mention that individuals in the chunking treatment group indicated that they actively trying to chunk together rhythms and pitch patterns, although “their overall performances suggested that they lacked motor skills to perform at a significantly higher level than their control-group peers” (p. 243). This quote makes the important point that chunking alone is not enough to promote expert behavior – it may be the case that music students who understand chunking principles may be unable to execute them effectively, given the fact that the cognitive structures required to benefit from chunking strategies, and
the fine motor coordination involved in music performance, likely develop at different rates within individuals. Ineffective execution of a chunking technique may also reflect task demands of a music performance, such as tempo (i.e., it is more challenging to correctly and technically execute a fast piece of music than it is a slow piece of music).

Across these studies, no overall benefit of chunking was observed. One potential reason for this may simply be due to the duration of the various interventions. While understandable given practical constraints, these interventions were quite short (Gaynor, 1995: four 50-minute sessions for two weeks; Hagen et al., 2012: 15 minute sessions, once per week for six weeks; Pike & Carter, 2010: 10 minute sessions, twice a week for three weeks). This is especially true in light of what is known about the amount of time required to form mental chunks (i.e., approximately 8 seconds for a chess player; Simon & Gilmartin, 1973), and the thousands of hours of practice required to achieve expert-level performance (e.g., deliberate practice: Ericsson et al., 1993). It may be that the variable of time (and specifically, the quality of time) plays a significant role in terms of chunk formation with practice; some authors advance the quality of time as one of most important factors in developing expertise (e.g., Ericsson et al., 1993). For this reason, it may not be surprising that brief academic chunking interventions are unsuccessful; while potentially impractical, the problem at hand begs for a proper longitudinal study.

One such study measured both expert and non-musician eye movements during a sight reading task over the course of a nine-month training period (Penttinen & Huovinen, 2011). These authors reported that non-musicians’ eye movements more closely approximated experts’ at the end of the intervention, as evidenced by longer saccades (i.e., larger melodic skips rather than reading note-by-note) (Penttinen & Huovinen, 2011). The authors interpret these results as evidence of skill development among the non-experts (i.e., evidence of the expert behavior of “perceptual grouping”). While it is not possible to say whether this is the result of certain instructional techniques, a longer intervention, or another variable, these results are nonetheless compelling. Motivation is also a critical
factor in expertise acquisition (for a related discussion, see McPherson & McCormick, 1999); the participants in Penttinen and Huovinen’s (2011) study were aspiring elementary school teachers. It may be the case that this group of participants was simply more motivated to complete the sight reading task than participants in other studies, resulting in significant improvement over time.

Related to the variable of time, one mechanism through which individuals may attain expertise is practice, or more specifically, *deliberate practice* (e.g., Ericsson et al., 1993). Deliberate practice can be defined as “a highly structured activity, the explicit goal of which is to improve performance” (Ericsson et al., 1993, p. 368). Deliberate practice is discussed with some frequency in relation with the acquisition of music reading expertise (for reviews, see Ericsson et al., 1993; Ericsson & Charness, 1994; Ericsson, 2002; Lehmann, 1997; Lehmann & Ericsson, 1997). Relatedly, other studies seem to suggest a “sensitive period” during which sight reading expertise must be acquired if it is intended to be mastered (i.e., before the age of 15) (e.g., Kopiez & Lee, 2006; 2008). This finding is especially compelling when considered in conjunction with the observation that informal practice does not improve sight reading skill (e.g., Zhukov, 2017). However, the degree to which a sensitive period may bear on theories of chunking remains unclear; it may be the case that chunk formation outside of this period is possible, but more effortful.

While deliberate practice represents a compelling mechanism by which experts hone their craft, an important question remains: To what degree is expertise acquired explicitly (as a result of the deliberate practice itself) or implicitly (as a result of the exposure afforded by deliberate practice)? While “quality of time” is an important facet of deliberate practice, so too are duration of time and amount of exposure to domain-related material. Research focused on disentangling the independent contributions of these different aspects of deliberate practice has the potential to help explain why there is often no observed benefit of instructional interventions using chunking principles (e.g., Pike & Carter, 2010).
Future research could also focus on disentangling the reason(s) why chunking interventions do not appear to improve sight reading ability. For example, one rather nuanced color-coding system of music notation that has yet to be tested may interest researchers studying musical chunking with the aid of color (Kuo & Chuang, 2013). This system has the ability to represent with color not just pitch height and duration, but also expressive markings and dynamics that may contribute more to musical chunk formation than is currently understood. It is also interesting to note that while extensive research exists on the eye movement mechanisms that contribute to learning to read text (e.g., Tunmer & Hoover, 2017), comparatively less work has addressed these questions about learning to read music. While this may be due to the fact that music reading is a difficult skill to acquire (possibly due to the lack of emphasis on reading music notation during instruction), more research on the acquisition of music reading skill is needed to understand this disconnect.
References


https://doi.org/10.3758/bf03194923


https://doi.org/10.3758/s13423-017-1406-5


https://doi.org/10.1117/1.601914


https://www.jstor.org/stable/40318992


https://doi.org/10.1561/1500000042


https://doi.org/10.1177/17470218211056924


https://doi.org/10.1007/978-1-4757-1968-0_18


https://doi.org/10.1177/87551233050240010106

https://doi.org/10.1109/milcom.2016.7795425

https://doi.org/10.1177/0305735618804038

https://doi.org/10.1080/00140139.2014.909952

*Music Education Research, 16*(1), 70–87.  
https://doi.org/10.1080/14613808.2013.819845

https://doi.org/10.1177/1029864916684193

https://doi.org/10.1177/0305735614550229
Table 1.

Means and standard error of the mean (in parentheses) for the variables analyzed in Experiment 1: Accuracy, Reaction Time (seconds), Fixation Duration (milliseconds), Number of Fixations, and Saccade Amplitude (degree of visual angle).

<table>
<thead>
<tr>
<th></th>
<th>Expert</th>
<th></th>
<th>Non-musician</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simple</td>
<td>Complex</td>
<td>Simple</td>
<td>Complex</td>
</tr>
<tr>
<td>Accuracy</td>
<td>.96 (0.01)</td>
<td>.96 (0.01)</td>
<td>.97 (0.01)</td>
<td>.90 (0.03)</td>
</tr>
<tr>
<td>Reaction Time (seconds)</td>
<td>5 (0.26)</td>
<td>15 (0.74)</td>
<td>9 (0.33)</td>
<td>23 (1.14)</td>
</tr>
<tr>
<td>Fixation Duration (ms)</td>
<td>329 (9.89)</td>
<td>388 (10.14)</td>
<td>385 (10.66)</td>
<td>444 (12.63)</td>
</tr>
<tr>
<td>Number of Fixations</td>
<td>16 (0.83)</td>
<td>37 (2.03)</td>
<td>22 (0.87)</td>
<td>51 (2.68)</td>
</tr>
<tr>
<td>Saccade Amplitude (*)</td>
<td>1.85 (0.04)</td>
<td>1.62 (0.03)</td>
<td>1.73 (0.04)</td>
<td>1.36 (0.04)</td>
</tr>
</tbody>
</table>
Table 2.
Means and standard error of the mean (in parentheses) for the variables analyzed in Experiment 2: Accuracy, Reaction Time (seconds), Fixation Duration (milliseconds), Number of Fixations, Saccade Amplitude (degree of visual angle), and OSPAN scores.

<table>
<thead>
<tr>
<th></th>
<th>Expert Upright</th>
<th>Expert Rotated</th>
<th>Non-musician Upright</th>
<th>Non-musician Rotated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>.90 (0.02)</td>
<td>.93 (0.02)</td>
<td>.84 (0.02)</td>
<td>.87 (0.02)</td>
</tr>
<tr>
<td>Reaction Time (seconds)</td>
<td>11 (0.71)</td>
<td>17 (0.77)</td>
<td>14 (0.74)</td>
<td>17 (0.79)</td>
</tr>
<tr>
<td>Fixation Duration (ms)</td>
<td>348 (11.74)</td>
<td>409 (13.37)</td>
<td>429 (12.20)</td>
<td>479 (16.26)</td>
</tr>
<tr>
<td>Number of Fixations</td>
<td>28 (1.88)</td>
<td>39 (2.04)</td>
<td>31 (1.66)</td>
<td>34 (1.82)</td>
</tr>
<tr>
<td>Saccade Amplitude (°)</td>
<td>1.57 (0.04)</td>
<td>1.39 (0.04)</td>
<td>1.48 (0.04)</td>
<td>1.35 (0.04)</td>
</tr>
<tr>
<td>OSPAN</td>
<td>0.67</td>
<td>0.71</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.

*OSPAN analysis. Panel A contains accuracy and OSPAN t-tests for Experiment 2; Panel B contains the bivariate correlations (** indicates significance at p < .01).*

A)

<table>
<thead>
<tr>
<th></th>
<th>Non-Musician</th>
<th>Expert</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>0.85</td>
<td>0.92</td>
<td>-2.41, p = 0.019</td>
</tr>
<tr>
<td>OSPAN</td>
<td>0.71</td>
<td>0.67</td>
<td>0.86, p = 0.394</td>
</tr>
</tbody>
</table>

B)

<table>
<thead>
<tr>
<th></th>
<th>1. Expertise</th>
<th>2. OSPAN</th>
<th>3. Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Expertise</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. OSPAN</td>
<td>-.11</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>3. Accuracy</td>
<td>.30**</td>
<td>.09</td>
<td>--</td>
</tr>
</tbody>
</table>
Figure 1.

An illustration of the foveal (1), parafoveal (2), and peripheral (3) regions of the visual field during the processing of a music score. The region shaded in gray indicates the location of the parafoveal region. Note: The size of the regions in this figure are approximate, and this figure is not drawn to scale.

Figure 2.

Example of potential chunks (green) and templates (yellow) in the chess domain (A) and the music-reading domain (B). In a chess board (A), a fork could be a potential chunk. A fork is a strategy in which a single piece (in this case, the White Bishop) is making at least two direct attacks simultaneously (the Black King and the Black Rook). The entire chess board could be a potential template. In a music score (B), a musical phrase could be a potential chunk, while the rest of the staff (or multiple staves) could be a potential template.

Figure reproduced from Maturi (2021), with author’s permission.
Figure 3.

Examples of music scores in the visually simple (Panel A) and visually complex (Panel B) conditions, and a schematic illustration of the trial progression in the music-related flicker paradigm (Panel C). The location of the screen change in Panel C is indicated by a red circle (shown here for illustration purposes).
Figure 4.

Means and standard errors for accuracy (Panel A), RT (Panel B), fixation duration (Panel C), fixation count (Panel D), and saccade amplitude (Panel E), as a function of expertise and complexity, in Experiment 1.

Figure 5.
Examples of music scores in the upright (Panel A) and rotated (Panel B) conditions, and a schematic illustration of the trial progression in the music-related flicker paradigm (Panel C). The location of the screen change in Panel C is indicated by a red circle (shown here for illustration purposes).
Figure 6.
Sample OSPAN trial from Experiment 2.

\[(6/2) + 5 = 9 \ ? \text{ Snake}\]
Figure 7.
Means and standard errors for accuracy (Panel A), RT (Panel B), fixation duration (Panel C), fixation count (Panel D), and saccade amplitude (Panel E), as a function of expertise and orientation, in Experiment 2.
Appendix A.

Demographic questionnaire used in Experiments 1 and 2.

Section 1: Basic Demographic Information

1. Sex: _________
2. Handedness: LEFT RIGHT AMBIDEXTROUS
3. Do you wear glasses or contact lenses? YES NO
4. Do you have normal (or corrected-to-normal) vision? YES NO
5. Age: _________

Section 2: Musical Background Information

6. Have you ever had any formal music training? YES NO
7. If YES, how many years?
8. If YES, what kind of setting did this training take place in (i.e., formal or informal)?
9. Do you still study music or play instruments? YES NO
10. If YES, in what ways do you still study music?
11. If YES, which instruments do you play, and how many years have you played each?
12. Are you able to read sheet music? YES NO
13. Have you had any formal or informal training in music theory? YES NO
14. Have there ever been any gaps in your music training or playing from when you started to the present time? YES NO
15. If YES, what is the size of the gap?
16. Have you ever participated in music training in an educational setting (i.e., classes in elementary/middle/high school)? YES NO
17. How would you rate your music theory knowledge?

   NO KNOWLEDGE   NOVICE   INTERMEDIATE   ADVANCED   EXPERT

18. How would you rate your ability to read sheet music?

   NO ABILITY   NOVICE   INTERMEDIATE   ADVANCED   EXPERT

Section 3: Musical Knowledge and Preference Information

For each statement, indicate the response that best fits with your personal feelings or beliefs about that statement.

Select from the following: Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree

19. I like to listen to music that speaks to my emotions.

20. When I find music that I like, I want to learn how to play and/or sing it.

21. I am fairly good at understanding the structure of a musical piece (e.g., time signature, key, tempo) just by listening to it.

22. I have a fairly good memory for music.

23. I would not consider myself a musician.