Cognitive modelling of early music reading skill acquisition for piano: A comparison of the Middle-C and Intervallic methods

Action editor: Hedderik van Rijn
Bruno Emond a,*, Gilles Comeau b

a National Research Council Canada, 1200 Montrea Rd., M50, Ottawa, ON, Canada K1A 0R6
b School of Music, University of Ottawa, 50 University, Ottawa, ON, Canada K1N 6N5

Available online 12 January 2013

Abstract

In the classical music tradition, knowing how to read music is an essential skill and is seen as a fundamental component to develop when learning to play the piano. This research’s focus is to study the possible impact of two different teaching approaches to the acquisition of initial reading skills. By using cognitive modeling, we are hoping to observe through computer simulation the problem solving and decision-making tasks involved in decoding a simple musical score. Our model intends to capture a novice initial coordination of music reading and motor operations on a piano keyboard. As such, it does not aim at modelling advanced sight-reading skills. The paper introduces the Middle-C and Intervallic methods followed by a description of an ACT-R cognitive model and simulation results upon learning with each of the reading methods. Inspection of the simulation results reveals differences in terms of declarative memory and cognitive processing demands. In particular, the Intervallic method requires a larger number of declarative knowledge related to notes, and more execution planning than the Middle-C method.

Keywords: Music reading; Middle-C method; Intervallic method; Computer simulation; ACT-R

1. Introduction

In the classical music tradition, knowing how to read music is an essential skill and is seen as a fundamental component to develop when learning to play the piano (Galyen, 2005; Sloboda, 2005). However, learning to read musical notation is a long and arduous undertaking (Anderson, 1981; Hahn, 1985) and, despite the value we attribute to it, it is not always successful. In North America and in Europe, piano book tutors are at the centre of a beginner student’s learning environment as piano teachers often rely on these books to provide the whole foundation of a pianist’s musical education and much of the initial training on reading musical notation (McPherson & Gabrielsson, 2002; Stewart et al., 2003). However, while having music reading as a common objective, the book tutors have introduced fundamentally different approaches such as the Middle-C, Intervallic or Multi-key approach; and more recently the Eclectic or Modified Multiple Key approach, which has supplanted the original Multi-key (Lomax, 1990). Surprisingly, despite the fact that the main focus of the piano tutors is the development of music reading skills, little is known about how this is done. Piano pedagogy textbooks provide long list of advantages and disadvantages for each of the different teaching approach (Uszler, Gordon, & Smith, 2000), however it is all based on intuition and on teachers experience and it has no experimental basis to support the analysis, or formal model of its development. Little scientific information is available to evaluate the real impact of each reading systems, to establish their efficacy and efficiency.
It is well recognized that there is a lack of cognitive models to explain how music reading is acquired. Hodges, the author of the Handbook of Music Psychology (1996) and author of a chapter on music reading in the Handbook of Research in Music Teaching and Learning (1992) wrote that "in music there is no theory devoted specifically to an explanation of music reading: thus, the bulk of the research appears to be devoid of a theoretical underpinning" (1992, p. 469). Sixteen years later, he confirmed that the situation was still the same (Lemay, 2008). The few theoretical models that have been proposed over the years are either still in an embryonic stages or entirely speculative and devoid of an experimental basis (Udtaisuk, 2005). The most well known cognitive model of music sight-reading was published by Wolf in 1976, and it was developed entirely based on interviews with four pianists (Wolf, 1976). It explains sight-reading as a problem-solving activity of pattern recognition, but no quantitative investigations were undertaken to refine and give legitimacy to the model. Fifteen years ago, Waters, Townsend, and Underwood (1998) realized a series of laboratory experimentation to observe how pattern recognition’s skills could play an important role in expertise musical sight reading and they have shown that in the pattern-recognition task, immediate recall of presented material correlate strongly with good sight-reading skills. Their study confirmed various experiments conducted previously by Sloboda (1978, 1985) to show the importance of pattern recognition in various tasks related to music reading. However, while pattern recognition seemed to be a promising avenue to help our understanding of music reading skills, Madell and Hébert (2008) deplore the fact that more recent trends in music reading research has been to experiment with the intricacy of eye tracking technology without a focus on pattern recognition (Kinsler and Carpenter, 1995). In addition, music reading studies deals with musicians who already know how to read music and have often reach the level of expertise. These models do not always shed lights on the skills required by a novice just being introduced to music notation. Without a solid model of music reading acquisition, it is not surprising that piano teaching material have come to propose very different approaches to music reading.

Piano playing is an elaborate skill that requires the coordination of many cognitive resources and subtle body movements. As such, expert piano playing performance has been the subject of many investigations (Altenmüller, Wiesendanger, & Kesselring, 2006; Hallam, Cross, & Thaut, 2009; Parnhurst & McPherson, 2002). However, the effect of pedagogical methods on novice performance and learning has not received the same level of attention from a cognitive point of view (McPherson, 2006). Empirical data on the effect of piano methods on learning are scarce, and very difficult to obtain in a controlled setting. As a first step to characterise the effect of pedagogical methods on novice performance and learning, a series of computer simulations were designed. The main objective of the simulations was to compare the resulting states of a common cognitive model after learning to play sequences of short piano pieces from different piano methods. The simulations focused on learning the association between the musical notation and the correct motor movements on the piano keyboard. The task to be performed by the model was a form of sight-reading task (Fourie, 2004). The task was to read a note on a music score, and play it on the piano. The model did not intend to capture looking ahead behaviour (Fourie, 2004), the representation and processing of musical sounds (Chikhaoui et al., 2009), learning motor skills (Jabusch, Alpers, Kopiez, Vauth, & Altenmüller, 2009), movement preparation (Palmer, 2005), and multitasking of music reading and motor movements as threaded cognitive tasks (Salvucci & Taatgen, 2008) were excluded from the models.

2. The Middle-C and Intervalic approaches

This research’s focus is to study the possible impact of different teaching approaches on the acquisition of initial reading skills. By using cognitive modeling, we are hoping to observe through computer simulation the problem solving and decision-making tasks involved in decoding a simple musical score. We want to examine how the different reading systems impact on the perceptual and motor processes. Since the Middle-C approach and the Intervalic approach have dominated the market for many decades now, we have selected two tutor series that are a good representation of each approach: The A.B.C. of Piano Playing: An Easy Method for Beginners (Berlin, Konicek, & Precious, rev. ed. 1983; original ed. 1941); The Music Tree: A Plan for Musical Growth at the Piano (Clark, Goss, & Holland, rev. ed. 2000; original ed. 1973; Clark first introduced the Intervalic approach under the title Time to Begin in 1955). These authors published their first tutor in the middle of the 20th century, both publications have gone through revision and re-edition and both are still in use by piano teachers. In order to understand the basic characteristics of the reading process involved in each approach, a quick overview of their reading system will be provided.

According to Lomax (1990), the Middle-C reading approach became influential in the early 1900s. Introduced by Mathews in Standard Graded Course of Studies for the Pianoforte in Ten Grades (1892), it was then popularised by the very successful tutors written by John Thompson Teaching Little Fingers to Play (1936) and the Modern Course for Piano (1936). Berlin’s A.B.C. of Piano Playing (1941) published a few years later and selected for our analysis was very much in line with the earlier Middle-C tutors. This reading approach requires the student to place the thumbs of each hand on middle C. The entire first piece is often played with that note only, and then on the following pieces, one note above and one note below middle C are introduced. As new notes are introduced, note names and traditional staff notation are learned simultaneously. The hand position with both thumbs sharing middle C and the other fingers resting on the surrounding white keys is
maintained generally for quite a long period of time so that
the student becomes familiar with these notes. This reading
approach was extremely influential throughout the second
half of the 20th century, Schaum and Cupp (1985) wrote that
“the Middle C approach continues to prevail because of
its unparalleled success and thoroughness. It is probably
the most widely accepted keyboard teaching system pres-
etly in use” (p. 68) and Lomax (1990) was affirming “the
Middle C Method is still one of the most widely used
approaches today” (p. 101).

In 1955, Frances Clark revolutionised the way that
music reading could be thought with the publication of
her Intervalic approach tutor *Time to Begin*. Elements of
this approach had been introduced earlier: partial-staff
notation in *Loomis’ Progressive Music Lessons* (Loomis,
1875) and the Landmark approach in *Year by Year Books*
(Williams, 1924). However, Clark was able to define the
Intervalic approach like no one had done before her and
she popularised it among piano teachers. She developed a
reading system where piano students are taught to read
music by recognizing intervals. As Uszler, Gordon, and
Mach (1991) explains “the Intervalic approach stressed
the development of spatial-directional reading habits con-
nected with the formation of hand-shapes and movements
that follow from intervalic recognition” (p. 107). Students
are encouraged to read by contour recognition and the
musical staff is introduced one line at a time. They are
thought to recognize steps (neighbouring keys) and skips
(skipping over one key) on a partial staff, then intervals
are introduced (seconds, thirds, fourths, etc.) and finally
they are given certain landmarks on the keyboard and they
are thought to distinguish the direction of the music
through intervals that are related to these guide posts.
Unlike the Middle C approach, the Intervalic approach
reinforces playing all over the keyboard.

3. Simulation of early music reading skills acquisition

This section presents the simulation methodology and
simulation results obtained by running an initial cognitive
model playing a series of musical staves belonging to either
the Middle-C or the Intervalic piano methods. The ACT-
R cognitive architecture was used to run the simulation
(Anderson et al., 2004). The simulation procedures con-
stisted of: (a) developing an initial cognitive model, (b) run-
ning the cognitive model with the different conditions
represented by the different sequence of music staves from
the two piano methods, and (c) comparing the model states
resulting from the separate simulations.

3.1. Initial cognitive model

The initial model contained only the minimal declarative
and procedural knowledge to be able to visually scan a
music staff for notes, the piano keyboard for keys, move
the hands and fingers over piano keys, press, hold and
release them, and the capabilities to process instructions
from a tutor. In addition to the content of the declarative
and procedural memories described in the following sec-
tions, the cognitive model also used base level activation
of declarative chunks, production rules compilation, and
reinforcement learning.

3.1.1. Declarative knowledge

The initial model assumed no prior knowledge of musical
notation, and of its association to specific key locations
on the piano keyboard. The only declarative knowledge
the initial model held were chunks about the association
between the number of beats (1–4), and the subjective per-
ception of time encoded as ticks. The model however had
chunks encoding the approximate duration of 1, 2, 3, and
4 beats (60 beats per minute) using the ACT-R temporal
module (Taatgen, Van Rijn, & Anderson, 2007; Van Rijn
& Taatgen, 2008).

Fig. 1a and b presents the visual encoding of the music
scores. As figure shows, both the Middle-C and the Inter-
valic methods share the same encoding, in spite of the dif-
fences in the layouts. The visual encoding of a note visual
location includes its X and Y absolute visual locations, its
relative horizontal and vertical visual locations, as well as
duration encodings using a combination of full or empty
circles, with or without stems, and with or without a dot.

Fig. 2 presents the visual encoding of the piano key-
board. This encoding is used to direct the hands towards
the proper key to associate with the encoding of the note
information on the music staff. The visual encoding of a
key location includes the absolute X and Y visual loca-
tions, the key color (black or white), the group type
(around 2 blacks or 3 blacks), the relative position of a
key in the group, as well as the relative position of the
group on the keyboard.

In addition to the visual encoding of the staves and the
keyboard, the model includes a chunk type representing the
knowledge about a note, which binds together the musical
notation information (staff, vertical location on the staff,
duration encoding), motor directives (number of beats,
hand, and finger to use), and associated key on the key-
board (group type, group position, key position in group,
and key colour). This representation aims at capturing
the visual characteristics of notes for musical notations,
and in this respect, it differs from a representation of its
sound properties (Chikhaoui et al., 2009).

Closely related to the note chunk, the model includes an
execution plan. An execution plan is basically a note chunk
augmented with the information about the horizontal pos-
tion of a note on the staff to encode the sequence of notes
to play, and the number of ticks (Taatgen et al., 2007; Van
Rijn & Taatgen, 2008) that the note should be pressed. The
execution plan acts as the control structure for the model’s
behaviour. Chunk slots are filled up based on visual
encoding and memory retrievals until the plan can be exe-
cuted. Plan execution chunks are held in the goal buffer of
the ACT-R cognitive architecture. The encoding for the
Fig. 1. ACT-R visual encoding of music staves.

(a) Middle-C

(b) Intervallic

Fig. 2. Visual encoding of the piano keyboard using ACT-R chunks.
note is similar to the theory of event coding where perception and action share a common representation (Hommel, Müßeler, Aschersleben, & Prinz, 2001).

3.1.2. Procedural knowledge

A total of 19 productions are part of the model’s initial procedural knowledge. These productions can be classified in productions for processing the tutor’s instructions (2), processing the visual information on the staff (2), determining the note duration (5), its key location on the keyboard (4), the finger and hand to use (4), and finally executing the motor action on the keyboard (2). Fig. 5 characterizes the overall flow of control in the model. The first task of the model is to attend the staff and encode the next note visual features. Then the model attempts to retrieve from declarative memory a note chunk using the visual features as cues. The retrieved note chunk slots are used (or guessed if no note is retrieved) to complete the missing information in the execution plan. The note duration, fingering and key location need to be determined in no particular order. Once the execution plan is completed, the model locates the key on the keyboard, move the hand and finger to the location, and press and hold the key for the given duration.

Fig. 3 also includes a description of the flow of control between the student model and an automated tutor. The tutor compares the note to be played by the student model to its performance and provide either a positive reward, or a negative reward with instructions. An instruction consists of a note chunk, correcting the note played. After the reception of an instruction, the model harvests its content to declarative memory, and proceeds to re-attend to same note on the staff. If the note played was correct, the model just proceeds to the next note on the staff.

3.2. Running the simulation

The simulation consisted of running a sequence of introductory piano pieces from the Middle-C method, and...
another one from the Intervalic method. For both sequences, the model started in an identical initial state (described in the previous section). Each sequence had 8 pieces and the model had to play every piece 5 times before moving to the next piece. The following pieces were used in the Middle-C and Intervalic conditions.

Middle-C (Berlin et al., 1983): second lesson right, second lesson left, third lesson right, third lesson left, fourth lesson right, fourth lesson left, sixth lesson right, sixth lesson left.

Intervalic (Clark et al., 2000): Take off, Landing, In a canoe, Space ship, Inchworm, Rock Band, On the bleachers, Halloween.

After each executed pieces, model states data were collected, in particular the number of declarative chunks in memory, as well as the trace of production rules execution, and their relative utility.

3.3. Results

Three types of data were collected during the simulation execution: the number of declarative chunks in memory, the trace of production rules execution, and their relative utility. The aggregated results are presented in Figs. 4 and 5.

Fig. 4 shows the number of declarative chunks in memory as the model progress through the execution of the 40 pieces of music (8 different pieces played 5 times). As the graphic shows, the Middle-C method (lower line) has a very gradual introduction of musical note information when compared to the Intervalic method. The main reason for this difference is somewhat obvious. Because the Intervalic method forces the learning musician to play over multiple octaves, the number of note chunks is therefore larger, reflecting the demands of the music scores.
Fig. 5 shows the percent of time spent by the model on building an execution plan, which means the exclusion of the time devoted to visual encoding and motor execution, and the inclusion of processes related to instruction encoding, retrieval, and filling up the execution planning chunk slots. A visual inspection of the graph seems to indicate that the Middle-C method (lower line) requires less retrieval and execution planning time than the Intervallic method. Similar to the previous result on the number of declarative chunks, the larger number of notes to be played with the Intervallic method demands more motor planning. However, the line threads seem to also have different patterns. The Intervallic method has more or less a constant planning time over the course of the simulation. On the other hand, the Middle-C method seems to require an increase of planning time. This increase could be correlated with the increase of notes in the method. The apparent consistency of planning time for the Intervallic method might reflect a ceiling effect cause by the constant number of features per note (location, duration, fingering).

Results from the production compilation indicated that the model learnt to skip productions, reflecting knowledge acquired about the meaning of the notes. Both methods generated similar productions and their utility values were comparable. For both piano methods, the utility values of new productions were larger than the initial production utilities, in particular for the productions related to the note information associated to the plan duration of a pressed keyboard note.

3.4. Discussion

The main objective of the simulation was to characterize the effect of different piano methods on the acquisition of piano playing skills as instantiated in a minimalist ACT-R cognitive model having just a few procedural knowledge units at the initial state. We focused on two observations from the simulation: (a) the number of declarative chunks being created, and (b) the time spent in planning before pressing a keyboard note. Our results indicated that the Intervallic method requires the creation of more declarative chunks, as well as more planning time than the Middle-C method. However, the main question remains as to the validity of the simulation as a model of novice piano learners in their interaction with piano methods. At this point in time though, only an indirect estimation of the value of the model is possible based on piano learning methods analysis by experts. We can also point to some empirical techniques that could serve to inform the cognitive modelling effort.

In her famous textbook in piano pedagogy, well-known piano pedagogue Uszler et al. (2000) is providing a comparative analysis of the advantages and disadvantages of the Middle-C approach and the Intervallic approach as it relates, amongst other things, to some elements of topographical awareness. Our simulation results seem to be congruent with some advantages of the Middle-C method. For example, Uszler, Gordon, and Smith analysis indicates that the Middle-C method provides: (a) easy visual guides on the staff and the keyboard, (b) a limited amount of pitch names and piano key locations, (c) a strong sense of the key of C before moving into other keys which fosters ear and hand security. These three features of the Middle-C method imply a smaller number of notes and keyboard keys to remember (less declarative chunks), and as a consequence, less planning required for mapping staff notes to keyboard keys. Both cognitive impact of the Middle-C method are reflected in the simulation results.

In contrast, some of the Intervallic method advantages point to the value of learning early a large portion of the keyboard, the simple staff notation focused on directional reading and pattern recognition rather than individual note naming and identification. According to our model, these perceived advantages introduce a larger number of declarative knowledge as well as more processing to perform keyboard operations as reflected in our simulation results. However, our minimalist model assumptions have a positive bias towards the Middle-C method in opposition to the Intervallic method. In order to capture the intended advantages of the Intervallic method, our model would have to be augmented with higher-level representations to capture notes to keys sequences, and musical phrases.

In spite of any improvement to the model though, it would have to be constrained by empirical data beyond expert analysis and predictions, and empirical techniques that could serve to inform the cognitive modelling effort need to be identified and explored. Eye tracking is still very experimental when it comes to music reading and it has been found that understanding eye movement patterns while reading music is much more complex then with words. Madell and Hébert (2008) have pointed out that in contrast to studies in text reading, research using eye movements to study music reading is “relatively undeveloped” (p. 157) and, so far, has produce very little results.

Researchers have used various methods of assessment to evaluate music reading: Eaton’s grading instructions (1978), Gilman’s scoring algorithm (2000), Gudmundsdottir’s error classification (2003) and Salis’ error categorization (1977) are all different methods that have been used to quantify errors done by performers while reading music; Lemay (2008) has adapted the Watkins-Farnum Performance Scale (for wind instrumentalists) for the context of piano performance; and more recently, Comeau (2010a, 2010b, 2009) has developed a tool to measure skills in music reading with a test comprising of original musical stimuli of progressive difficulties, a system of coding errors, and a grid to score and assess the music reading performance of pianists, both novice and advanced. Other forms of data collection grounded in music education and assessment could use evaluation by a jury of expert. This technique is still the principal type of music reading assessment used in examinations and competitions (Lemay, 2008) and many studies focussing on the cognitive processes involved in music reading have used expert evaluations to rate reading performances (Furneaux & Land,
4. Conclusion

Advanced music reading skills exhibits a smooth coordination of visual encoding and motor skills (Fourie, 2004; Kopiez & Lee, 2008). With skill development, this combination requires a transition from multitasking to cognitive processes concurrency. As notes are being read on the staff, motor movements are planned and executed, while the reading process is progressing beyond what is currently played. Reading efficiency demands the coordination of psycho-motor speed, early acquired expertise, mental speed, and the ability for auditory imagery (Kopiez & Lee, 2008).

As an initial step towards characterizing the effect of different piano methods on the acquisition of piano playing skills, we constructed a minimal cognitive model which acquired declarative and procedural knowledge through the execution of novice piano pieces form the Middle-C and Intervallic methods. Our model intended to capture a novice initial coordination of music reading and motor operations on a piano keyboard. As such, it did not aim at modelling advanced sight-reading skills (Dirkse, 2009) like topographical awareness of the keyboard (limited need for visual feedback to find keys), fluency in directional reading (notes spatial relationship over individual note identification), pattern recognition abilities (grouping in musical phrases), and habits of effective sight-reading execution. Inspection of the simulation results revealed differences in terms of declarative memory and cognitive processing demands. In particular, the Intervallic method requires a larger number of declarative knowledge related to notes, and more gesture planning than the Middle-C method.

There are some limitations to the current state of the research that need to be mentioned. In particular the model would need to integrate a representation of sound to a note (Chikhaoui et al., 2009). This is important because the inner playing of a piece of music is a good determinant of music reading performance (Fourie, 2004). Also the model only focuses on individual note and has no notion of musical patterns. A more realistic model of motor movement could also be added, but mostly the model should be able to adress the visual and motor concurrency and the development of reading ahead strategies. The model does not aim at modelling errors. For example Fourie (2004) reports that 80% of error in sight-reading are rhythmic in nature, probably caused by the difficulty related to locating the correct key on the keyboard. This measure could be an interesting one in comparing the Middle-C and Intervallic methods, given the larger number of keyboard keys in the latter method. In this respect, the model should also have a representation of intervals, which at the moment is not present. Note accents were left out of the simulation, even though it is present in the introductory pieces of both the Middle-C and Intervallic methods.

Acknowledgments

We would like to thank the Social Sciences and Humanities Research Council for their financial support.

References


