BetterNote: Evaluating Alternative Music Notation

Jack Swiggett
Stanford University
610 Mayfield Ave, Stanford, CA 94305
swiggett@stanford.edu

ABSTRACT
This paper presents a tool for objectively evaluating music notation systems. In particular, we assess the ease with which a musician playing piano can sight-read music written using a particular system. We use this tool to evaluate two different notation systems—traditional Western notation and a proposed alternative, known as *Klavarskribo*. We present the results of this evaluation, and discuss future steps that would further our understanding of alternative music notation systems.

Author Keywords
Music, notation, piano, usability, user study.

ACM Classification Keywords
H5.2. Information interfaces and presentation (e.g., HCI): User Interfaces.

INTRODUCTION

Western music notation has a long and complex history. It began with systems of symbols used to aid singers during the 9th to 12th centuries. Over time, these symbols were mutated and augmented to represent a greater variety of pitches, rhythms, and stylistic choices. Notation arrived at its present form during the 17th and 18th centuries, and since then composers have continued to add and modify symbols, but the fundamental principles have remained the same [1].

Despite its haphazard history, Western music notation is ubiquitous today. It is used across all instruments and most musical styles. With few exceptions, accomplished musicians are intimately familiar with it. Yet it is not an easy system to learn. It relies on a five-line staff (Figure 1), in which musical pitches are mapped to different lines and spaces. This mapping can be confusing and difficult to learn, made worse by the fact that a staff can have one of several different "clefs," each of which implies a different pitch mapping. And the lines and spaces are not enough to represent every note in Western music theory, so notes are sometimes augmented with a "sharp" (♯) or a "flat" (♭) to change their pitch. On top of this, learners must study a separate system for representing rhythm.

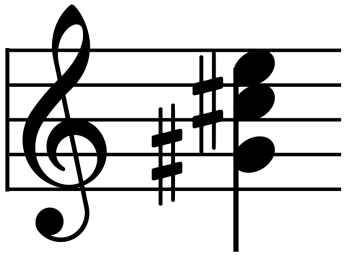


Figure 1. A three-note chord on a staff in Western notation

As a result of these complexities, beginning musicians are often intimidated by notation, and many never learn to read it. Yet without reading notation, they cannot easily share their musical ideas or explore the vast wealth of written music that exists. Many have sought to remedy this problem by developing alternative notation systems that are more intuitive for learners. Yet while hundreds of alternative systems have been proposed, we have almost no objective evaluations of their efficacy. Without a way to compare systems, we cannot redesign and optimize music notation to serve a wider audience.

Klavarskribo

Our research focuses on one particular alternative notation, called Klavarskribo. Klavarskribo was invented in 1931 by Cornelis Pot of the Netherlands [4]. The lines of the staff are oriented vertically, and music is read from top to bottom rather than from left to right. The lines of the staff correspond to the black keys of a piano, and the spaces to the white keys (Figure 2), making it especially suitable for individuals learning piano.

In this paper we present *BetterNote*, a software tool for evaluating the ease with which someone playing piano can sight-read a certain type of music notation. We present a methodology to evaluate both Western music notation and

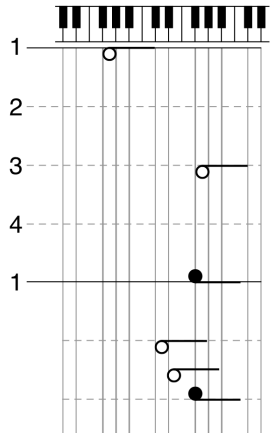


Figure 2. A short melody in Klavarskribo notation, with a piano keyboard shown for reference

Klavarskribo using this tool. We then discuss our results from performing this evaluation on ten test participants, with varying levels of musical experience. Finally, we discuss next steps to improve our understanding of alternative music notation systems.

RELATED WORK

Existing research surrounding music notation is limited. The Music Notation Modernization Association (MNMA) conducted one notable study in the 1990s, looking at over 500 notation systems. Trained musicians evaluated the notation systems using a combination of objective and subjective criteria, narrowing them down to the two highest-rated systems [2].

While this study is interesting, it has many flaws. Since notation systems were evaluated by experienced musicians, the evaluators were likely to be biased towards systems that were similar to the Western notation they had been reading all their lives. In addition, the screening criteria included strict rules that seem to be focused less on usability, and more on the goals of the study's authors. For example, one criterion mandated a "fully proportional pitch coordinate, where … progressively larger pitch intervals have progressively larger spacing on the coordinate, providing a visual representation of each interval that is *exactly proportional* to its actual sound [emphasis added]." Both Western music notation and Klavarskribo show pitch in an approximately, but *not* fully proportional way, so would fail this criterion, yet both have proven successful as notation systems.

Researchers have also looked at adding color to notation to facilitate learning. A study published in 1991 looked at colored music notation as an aid for fifth- and sixth- grade wind players learning to read traditional notation, with different pitches highlighted in different colors [3]. The study found that students who learned using color-coded notation strongly preferred it to black and white notation. However, when those students had to then read black and white notation, they were significantly worse at sight-reading than classmates initially trained on black and white notation. Since the goal of the study was to use color as a stepping-stone towards reading traditional notation, the technique was deemed ineffective.

While color was not effective in teaching traditional notation, it might be a useful tool in an alternative notation system. However, it is less convenient to write and print, and might pose a difficulty to individuals with color blindness. We have therefore chosen not to explore it here.

SOFTWARE DESIGN

BetterNote is a web application designed to run in a web browser. It is built in JavaScript, using React.js for the client interface, Node.js for the server, and MongoDB to store persistent data[1].

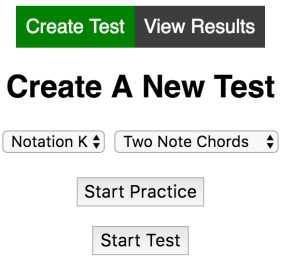


Figure 3. The BetterNote "Create Test" page

Creating a Test

The "Create Test" page (Figure 3) allows a user to launch a new notation test. She selects the notation type: "Notation K" for Klavarskribo or "Notation T" for traditional notation. She also selects the types of chords to be included in the test: "Single Notes," "Two Note Chords," or "Three Note Chords." Then she can either start a practice session, or start a test.

Executing a Test

In both cases, the test executes as shown in Figure 4. A piece of music notation is shown at the top of the page, representing a note or chord using the selected notation system. A piano keyboard is shown below. The user must select the keys on the keyboard that correspond to the notation by clicking on them. Selected notes are highlighted in orange. Once the necessary number of notes has been selected, the "Next" button is enabled, and the user clicks it to move on to the next chord. There is no way to deselect a note once it has been clicked, just as musicians are not able to go backwards and retry something when sight-reading a piece of music (at least when playing with others).

In a practice session, once the required number of notes has been clicked, the correct notes are marked in green (Figure 5). This lets the tester see if they are understanding the notation correctly, and correct any errors in their understanding so they can accurately identify the subsequent chords.

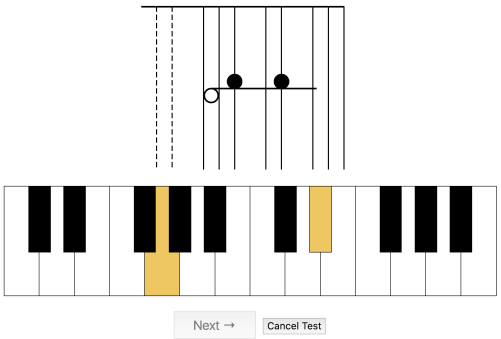


Figure 4. The BetterNote test interface

[1] BetterNote source code and installation instructions are available at https://github.com/jackswiggett/BetterNote.

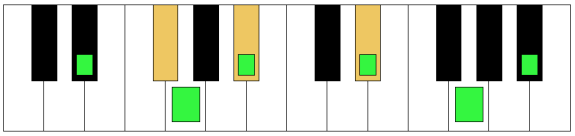


Figure 5. The keyboard in a practice session, after all three notes have been identified. Correct notes are marked in green. We see that the tester identified two of three notes correctly.

In a test session, the correct notes are not marked. The user simply moves on to the next chord and continues until the test is complete. The application records data about each chord, including the fraction of notes that were identified correctly and the time required to identify all the notes in the chord. Accuracy is calculated without regard to the octave in which the note was selected (for example, if the notation represents a D#, the tester can select either of the two D# keys that appear in the virtual keyboard). Once the test is complete, a detailed test log is sent to the database.

Viewing Test Results

The "View Results" page (Figure 6) loads and displays the test logs stored in the database. Test logs can be searched by the name of the tester. They can also be filtered by notation style, and the number of notes per chord in the test. Logs are filtered and redisplayed as the user types, so that it is easy to find a specific test log.

For each test log, BetterNote calculates two summary statistics. The first, "overall accuracy," is the total number of notes that the tester identified *correctly* during the test, divided by the total number of notes in the test. It can also be thought of as the average per-chord accuracy—so if a tester identifying three-note chords identified, on average, two of the three notes correctly, their overall accuracy for the test would be 67%.

The second statistic is called "median elapsed time." An elapsed time is calculated for each chord in the test, as the number of seconds between when the notation is first displayed to the tester, and when the tester has identified (correctly or incorrectly) all the notes in the chord. We

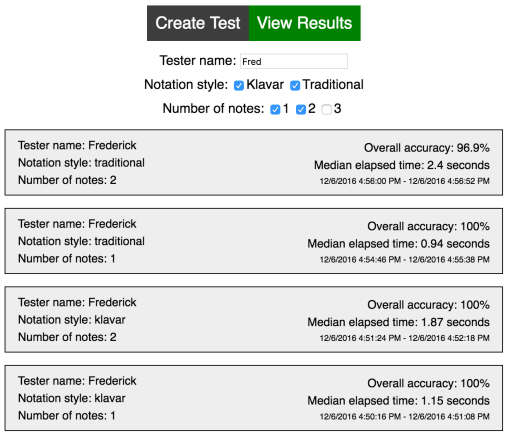


Figure 6. The BetterNote "View Results" page

calculate the median elapsed time among all the chords that the tester identified during a test. We use median rather than average to avoid an inaccurate skew towards longer times. Especially among inexperienced testers, we expect that a tester might pause for a particularly long time on a chord because they have forgotten or become confused about an aspect of the notation, and need time to remember how it works. The average would be heavily influenced by these anomalies, while the median effectively ignores them.

RESEARCH METHODOLOGY

This paper presents research on ten test participants. The participants were a convenience sample of individuals who were available to do the test. However, the sampling method should not significantly impact the significance of the results, since we do not seek to measure music reading competency among the general public, only to compare two different types of notation among the same participants.

For each participant, the experiment was conducted as follows:

1. Ask the following two questions, and record the participant's responses:
   - How much experience do you have reading traditional music notation?
     - None
     - A little bit
     - A moderate amount
     - A lot
   - How much experience do you have playing piano?
     - None
     - A little bit (< 1 year)
     - A moderate amount (1 - 3 years)
     - A lot (> 3 years)
2. Have the participant read a brief document describing what they will be doing during the test. This document is provided in Appendix A.
3. Randomly select either traditional notation or Klavarskribo as a starting notation. Then, for the selected notation:
   a. Provide the participant with a document describing the notation, and how to use it to identify notes on a piano keyboard. The documents used to teach the notations are provided in Appendix B. We have designed the documents to use roughly the same style of writing and imagery, and to provide roughly the same level of guidance for each notation.
   b. Tell the participant that they have up to 15 minutes to study the document, and to ask the person conducting the test any clarifying questions. They may also do as many practice tests as they like using the BetterNote interface.

Test Type	Notation Type	Avg (s)	Std Dev (s)
Single Note	Traditional	1.49	1.14
	Klavarskribo	1.10	0.41
Two Notes	Traditional	7.04	6.03
	Klavarskribo	2.41	0.74
Three Notes	Traditional	5.90	7.03
	Klavarskribo	3.87	1.24

Table 1. Average and standard deviation of the "median elapsed time" required to identify a note or chord, across the ten test participants.

c. Once the participant is ready, or 15 minutes have passed, have them do the three tests in BetterNote using the given notation, in the following order: single notes, two-note chords, three-note chords. The three tests are designed as follows:

   Single notes: 24 randomly selected notes, encompassing all the notes of the scale.

   Two-note chords: 16 randomly selected chords, encompassing intervals from a minor second to a major seventh.

   Three-note chords: 12 randomly selected chords, constricted to major and minor triads in all three inversions.

   Each test is designed to terminate once the tester has identified all the chords, or when 2.5 minutes have passed. This ensures that slower testers will not spend an inordinate amount of time completing the test. Since elapsed time is calculated on a per-chord basis, this does not negatively influence the accuracy of our results.

4. Repeat steps 3a-c for the other music notation system.

RESULTS

Each participant completed a total of six tests—three for each style of music notation. On average, participants were very accurate at identifying notes. The average accuracy across all the tests conducted was 97.5%, with a standard deviation of 4.3%. We saw much greater variability in the median time required to identify the notes (Table 1), especially when testing on traditional notation. In that case, the standard deviations were roughly as large as the averages for all three tests, indicating a wide range of test results.

Because of the relative homogeneity of the accuracy measurements, it would be difficult to use them to compare the two notations. We therefore do not analyze our data separately for the two different metrics (time and accuracy). Instead, we calculate a unified score for each test, using the following formula:

$$\frac{\textit{fraction of notes identified correctly}}{\textit{median time (s) required to identify a note/chord}}$$

This formula gives a higher score to testers who identified notes more accurately, and a lower score to testers who took a longer time to identify the notes. For each tester, we then calculate an overall score for traditional notation and for Klavarskribo, by averaging their scores across the three tests.

For each tester, we also calculate a metric for their prior experience, based on their responses to the two initial questions (step 1 of the research methodology). For each response, a score is calculated as follows: none = 1, a little = 2, a moderate amount = 3, a lot = 4. We average the scores for the two questions to get an overall metric.

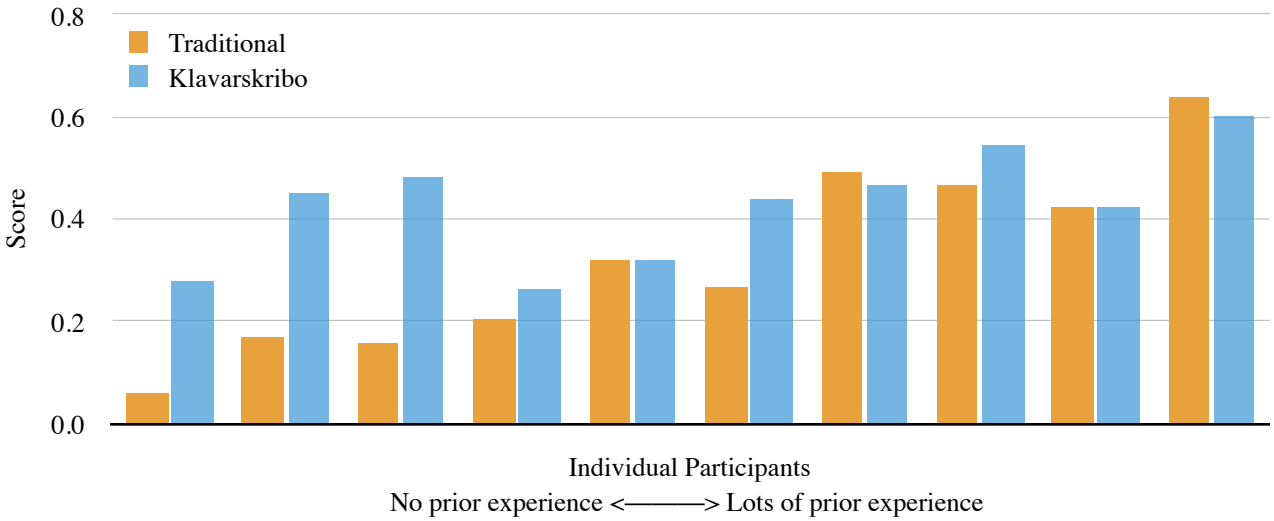


Figure 7. Individual participants' scores for traditional and Klavarskribo notation. Participants are sorted by prior experience, as indicated by the label below the x-axis.

Figure 7 shows the test scores for each of the individual testers. The testers are sorted by their prior experience metrics, so individuals with greater prior experience appear further to the right. Scores for traditional notation are shown in orange, and scores for Klavarskribo in blue. We can see that among the three leftmost testers (those with the least prior experience), scores for Klavarskribo notation are much higher than those for traditional notation. Among the four most experienced testers, however, the scores for both notations seem to be roughly the same.

In Figure 8, we divide the test participants into two groups based on their prior experience metrics. Those with a metric less than 3 are considered to have "none" or "a little bit" of prior experience. Those with a metric of 3 or above are considered to have "a moderate amount" or "a lot" of prior experience. Four participants fall into the first category, with the remaining six participants in the second category. We calculate average scores for the two groups of participants, for both notation styles.

Error bars show a 95% confidence interval around the mean score for each group. Our confidence interval calculations may be inaccurate at such a small sample size, and may not apply to the general population given that we do not have a random sample. However, setting aside these limitations, there does appear to be a statistically significant difference between the performance of inexperienced participants using the two notations. Participants with little to no experience appear to have a significantly easier time reading Klavarskribo notation than reading traditional notation. Among experienced participants, we do not notice a statistically significant difference between the two notations (although one might appear if we had a larger sample size).

Looking at just traditional notation, we see that experienced participants did significantly better on the tests than inexperienced participants. When using Klavarskribo, however, both experienced and inexperienced participants performed roughly equally well. Again, we might observe a difference with more participants, and the utility of these results is limited by the inability to collect a random sample.

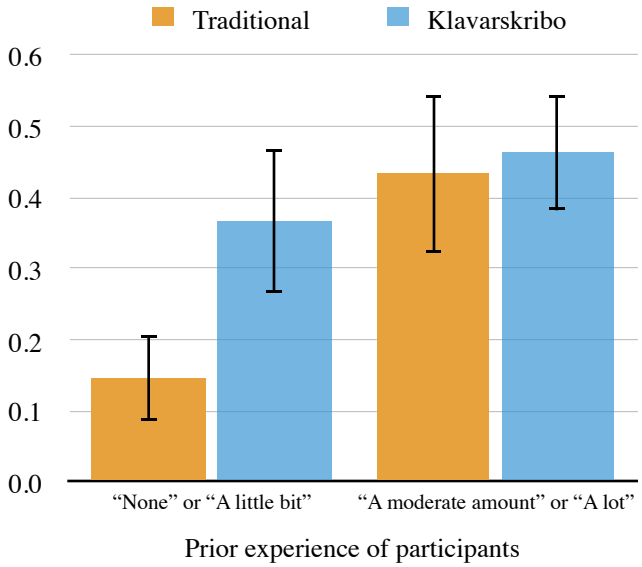


Figure 8. Average participant scores, shown separately for participants with limited prior experience (N = 4) and participants with significant prior experience (N = 6). Error bars show a 95% confidence interval.

DISCUSSION

We can draw several conclusions from these results. Most importantly, Klavarskribo seems to be significantly easier to learn for inexperienced participants. It could therefore be a valuable tool for non-musicians, or musicians who have not learned to read music. People who intend to pursue music seriously or professionally may still choose to learn traditional notation, given its near-universal use. However, many individuals may simply want to learn a few of their favorite songs, or record some ideas they've come up with on the piano. For these individuals, learning the particularities of traditional notation may seem like too much effort, with too little return. Klavarskribo could provide a way for such individuals to better enjoy music, and explore the wide variety of written music that exists.

We also observe that even among experienced musicians, Klavarskribo seems to be roughly as easy to read as traditional notation, after less than fifteen minutes of training. Experienced musicians generally spend weeks learning to read traditional notation, and months or years before they become highly proficient. This means Klavarskribo may be much quicker to learn than traditional notation. Importantly, this also means that it is not terribly difficult to adapt to Klavarskribo, even if the musician is used to a different notation system. If Klavarskribo were to see wider adoption—for example, among orchestras or in other contexts with experienced musicians—those who had learned to read traditional notation would likely be able to adapt to Klavarskribo without too much difficulty.

Limitations

These results have several important limitations. We have already discussed the limitations imposed by the sample used for the study. Study participants also responded in different ways to the instructions, potentially impacting results. Some participants used a large portion of the time allotted for practicing a given notation. Others, however, decided to move on to the test component of the study as soon as they felt like they had grasped the basics of the notation, spending very little time practicing. We were also limited by the fact that we had to present one notation before the other. Participants may have gone more quickly or paid less attention to the second notation because they wanted to finish the study quickly, or were tired of completing the tests. At the same time, participants were more familiar with the BetterNote system when testing with

the second notation, which may have improved their scores. Well we attempted to account for this by randomizing the order of the notations, we would need a larger sample to ensure that this did not impact our results.

We were also limited by the need to teach absolute beginners two notation systems in a reasonably short time period. We could not include longer pieces of music with many notes and chords, and we could not teach systems for representing rhythm, since these would have taken too long to learn. However, such situations are much more common in real-world sight-reading than simply reading a single note or chord.

Finally, BetterNote provides a virtual keyboard, which is not the same as a real piano. When playing a real piano, the musician must move her entire arm and hand, rather than just clicking a mouse. In addition, chords should be played so that all of the notes are struck simultaneously, something which is not possible when clicking with a mouse. We were unable to realistically simulate these conditions for this study.

The next section discusses some solutions to these problems, as well as other future directions for research.

FUTURE WORK

A simple next step would be to perform the same evaluation using BetterNote, with a larger sample size and/or with additional alternative notation systems. Klavarskribo is only one alternative among many, and different alternatives are likely to have different benefits and drawbacks. BetterNote could also be extended to support input via a MIDI keyboard. This would better simulate the experience of playing a piano, leading to more accurate results.

A second step would be to conduct a similar study with beginning musicians over a longer time period, teaching them to read longer, more complicated music. Such a study would assess reading pieces more similar to those that musicians would realistically want to learn, and would incorporate assessments of accuracy, rhythm, musical style, and ease of use (for example, ease turning pages). It could require anywhere from several hours to several weeks (or longer), depending on the scope of the results desired.

Another study might teach Klavarskribo or another notation to performing musicians who have significant experience with traditional notation, and have them read complicated music in both notations. It would assess whether musicians can achieve the same competency reading Klavarskribo as they have reading traditional notation, and how long this takes. It might also assess differences in the resultant understanding of the music. For example, does one notation better support memorization of music than other?

BetterNote is limited to sight-reading for pianists. While piano is often an instrument of choice for beginning musicians, a notation system should be effective for all instruments. One import area of research consists in extending this methodology to other instruments, and developing comparable tools to evaluate reading ability. For example, a software tool might listen to a musician via a microphone, and analyze the pitch of the notes played to determine their tonal and rhythmic accuracy. Such a system could be applied to a wide range of instruments, including voice.

Future research might also look at uses of music notation other than sight-reading. For example, is one notation easier to read for an orchestra or band conductor? Do composers find it easier to quickly transcribe their ideas in one notation than in another? Is a certain style of notation harder to sight-read, but more effective in a performance setting when the musician is already familiar with the piece?

CONCLUSION

Music is integral to many people's lives, yet few can read music notation as fluently as they read written words. Alternative music notation systems have the potential to make music notation more widely accessible, and enable more people to play and enjoy music. Yet to select and improve upon these systems, we must develop methodologies to objectively evaluate them. In this study, we found that in certain contexts, Klavarskribo presents distinct advantages over traditional notation for beginning learners. However, we have only begun to explore the potential of detaching ourselves from Western music notation. We should continue examining new ways to represent music, and to democratize our ability to transcribe and share it.

REFERENCES

1. Bent, I.D. (2007). Musical notation. *Encyclopædia Britannica*. https://www.britannica.com/art/musical-notation.
2. Research Project. *Music Notation Modernization Association (MNMA)*. http://musicnotation.org/mnma/research-project/.
3. Rogers, G. (1991). Effect of Color-Coded Notation on Music Achievement of Elementary Instrumental Students. *Journal of Research in Music Education*, 39(1), 64-73. http://www.jstor.org/stable/3344609.
4. The Klavar Method. *The Klavar Music Foundation*. http://www.klavarmusic.org/Explanation.pdf.

Appendix A. Introduction given to test participants

Introduction

This guide will teach you the basics of two different music notation systems. A music notation system is a way of using symbols to represent the notes that a musician should play on an instrument. For the purposes of this guide, you are the musician, and your instrument is a piano.

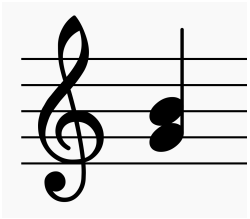
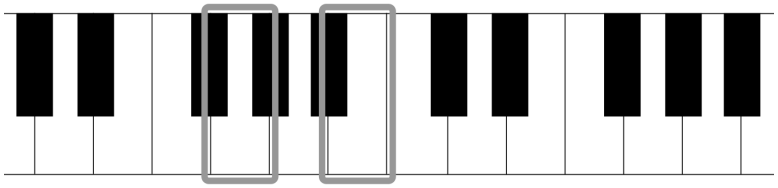
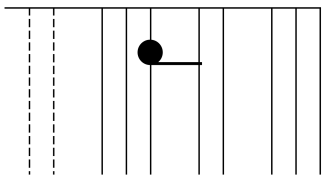
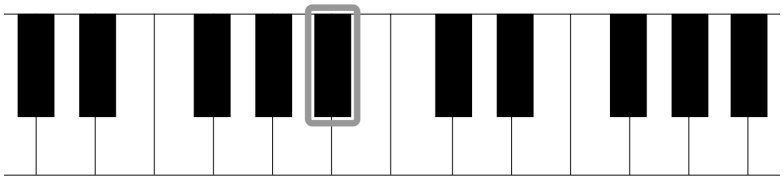
You will learn to:

- Read a piece of music notation that represents a single note on a piano, then identify that note on a virtual piano.
- Read a piece of music notation that represents a chord (two or more notes played at the same time), then identify that chord on a virtual piano.

You do not need to worry about reading rhythm.

Examples

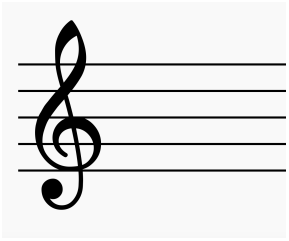
The following examples give you an idea of what you will learn to do. At this point you are *not* expected to understand how the music notation works; you will learn to read it later. These two examples use two different systems of music notation. You will learn to read both systems.

You read:	You identify these keys on a piano:
	
	

Appendix B. Documents given to test participants, describing the two notation systems

Music Notation System T

This notation system uses five horizontal lines, as well as the spaces between those lines, to represent the keys of a piano. These lines and spaces are called the *staff*.

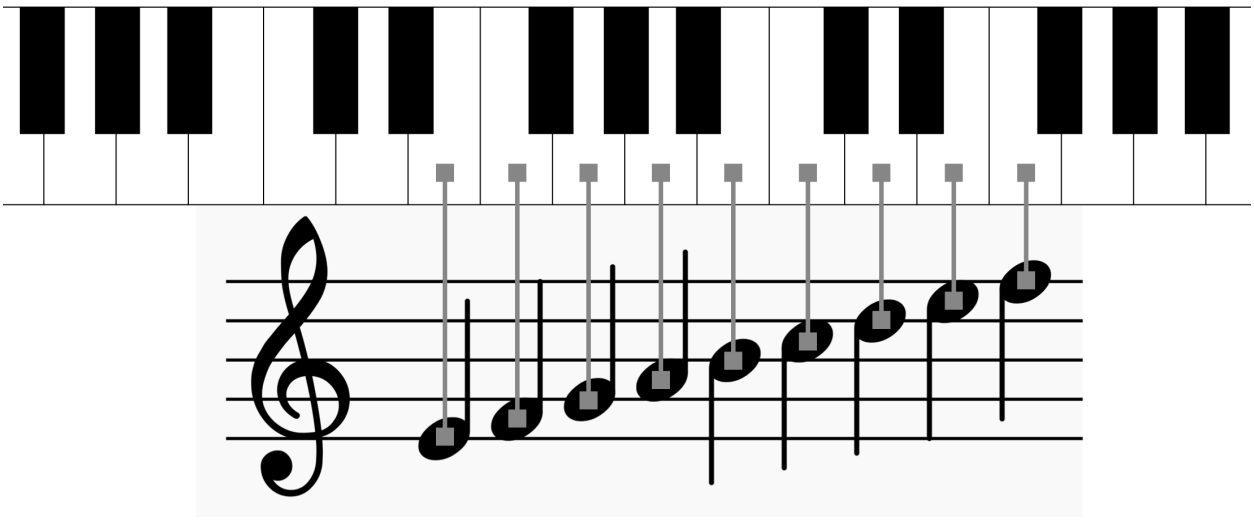


The staff

> *Note: Don't worry about the big, curly design at the left of the staff. It is part of this notation system, but you may ignore it.*

Every line and space between lines on the staff corresponds to a different white key on the piano. Lines and spaces that are lower on the staff correspond to keys that are further left on the piano. A musical note, represented by a black oval, is placed on one of the lines, or in the space between two lines. This indicates that you should play the corresponding key on the piano.

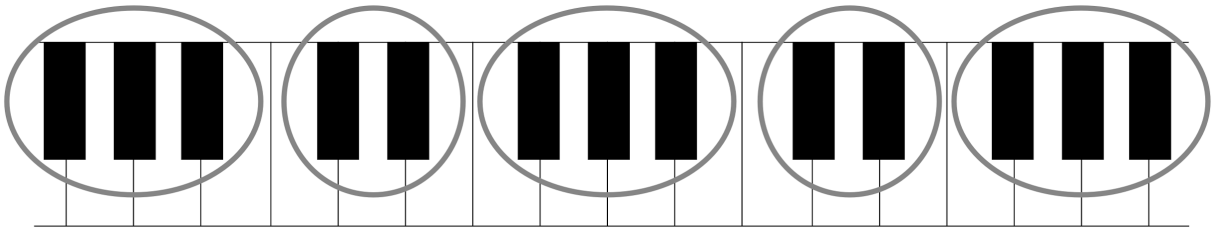
The following diagram shows several notes on the staff, and the corresponding white keys on the piano. Don't worry about black keys for now.



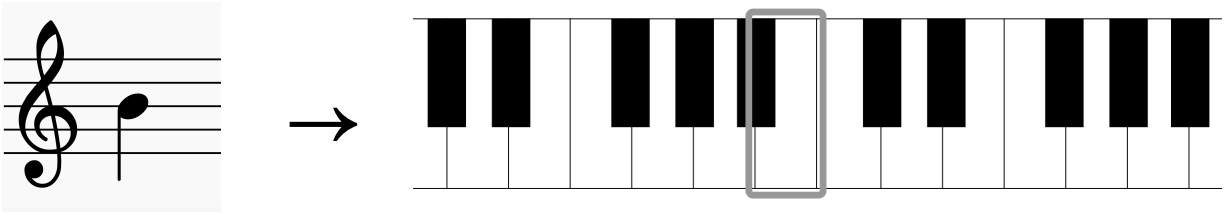
> *Note: Don't worry about the black lines extending upwards or downwards from the notes. These are part of the notation system, but you may ignore them.*

Finding notes

You will have to look at a note on the staff, and find the corresponding key on the piano. How can you find the correct key without having to reference the above diagram? First, look at the layout of the piano keys. The black keys on the piano come in alternating groups of two and three:



The *middle line* of the staff corresponds to a white key that is *just right of a group of three black keys*:



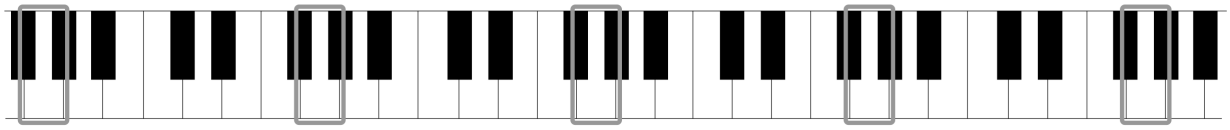
Given a note on the staff, you can find the corresponding white key by counting up or down from the middle line of the staff, and counting the same number of steps right or left from the white key indicated above. Note that:

- Lines and spaces on the staff are *equal size* steps on the piano; each line and space corresponds to a single white key.
- Moving from one line to the next line on the staff, or one space to the next space, is the same as moving by two keys on the piano (i.e. skipping a single white key).

Do not worry about *which* group of three black keys you use as a point of reference. For the purposes of this guide, you can use whichever one you want (high or low on the piano). So if you read

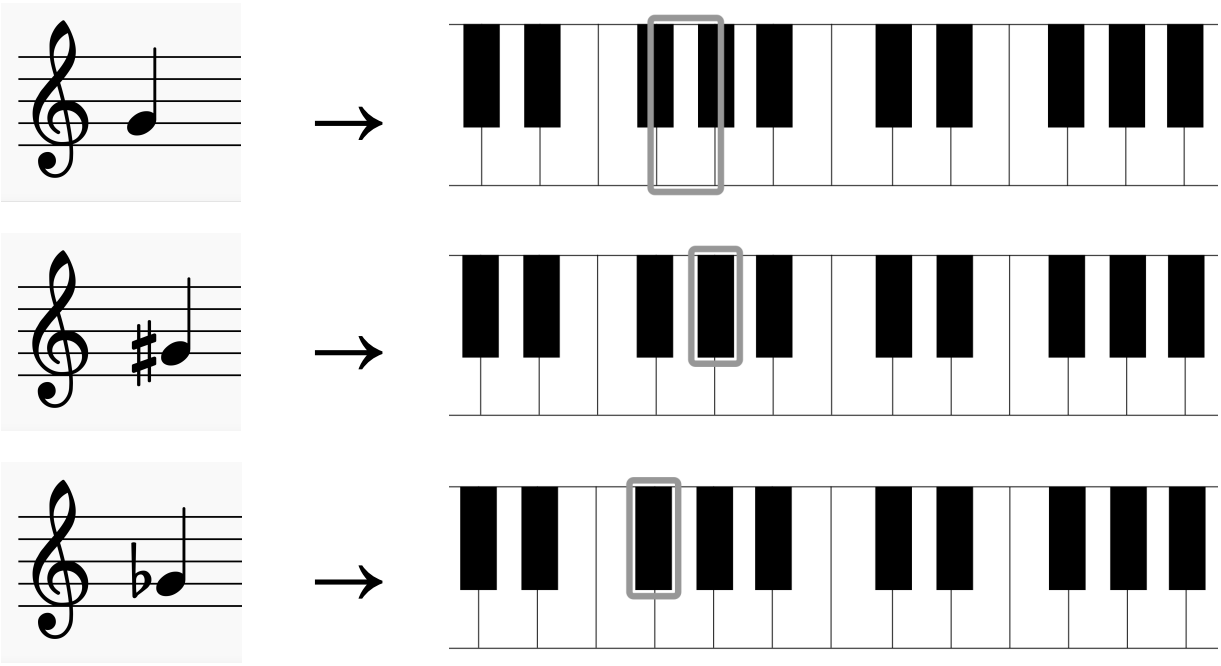
,

you can play any of the following keys:



Playing black keys

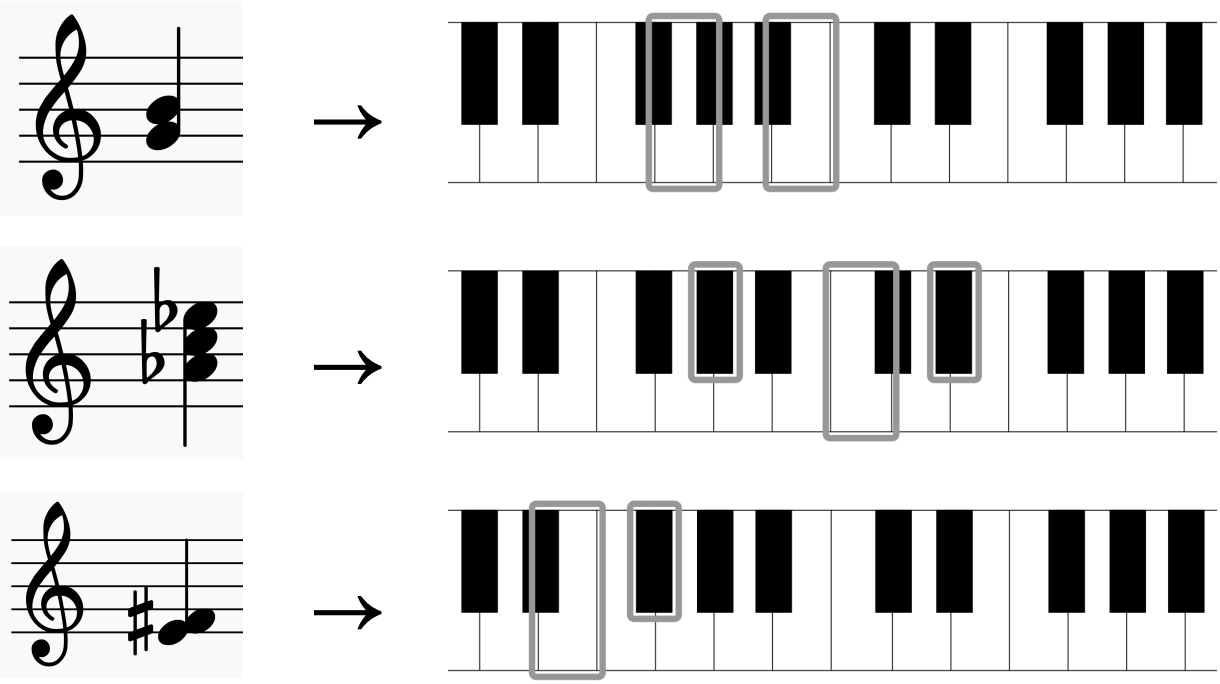
Black keys are represented on the staff in the same way as white keys, but adding either a *sharp* symbol (♯) or a *flat* symbol (♭). A sharp symbol is used to indicate the black key that is directly *above* the given white key. A flat symbol is used to indicate the black key that is directly *below* the given white key. For example:



Note that every black key can be represented in two ways — either as the white key below it, with a sharp symbol; or as the white key above it, with a flat symbol.

Playing chords

A chord is two or more notes, played at the same time. Notes in chords are represented the same way as single notes. The notes are simply stacked on top of each other:



Observe two things in the last example above:

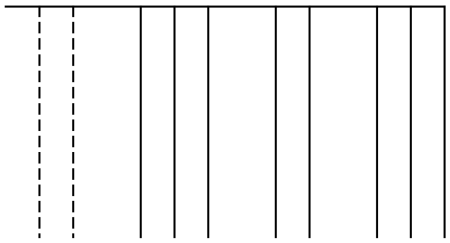
- The two notes on the staff are slightly horizontally offset from each other. This is simply to prevent them from overlapping, since they are so close to each other.
- The sharp symbol is applied to the note *directly* to the right of it on the staff. In this case, it is applied to the note that is in the space between two lines, not the note that is on a line.

\- - -

This concludes the introduction to this music notation system.

Music Notation System K

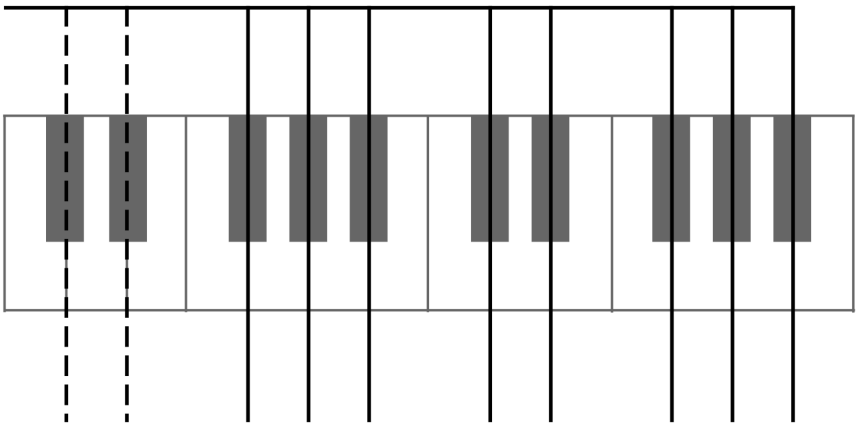
This notation system uses vertical lines, as well as the spaces between those lines, to represent the keys of a keyboard. These lines and spaces are called the *staff*.



The staff

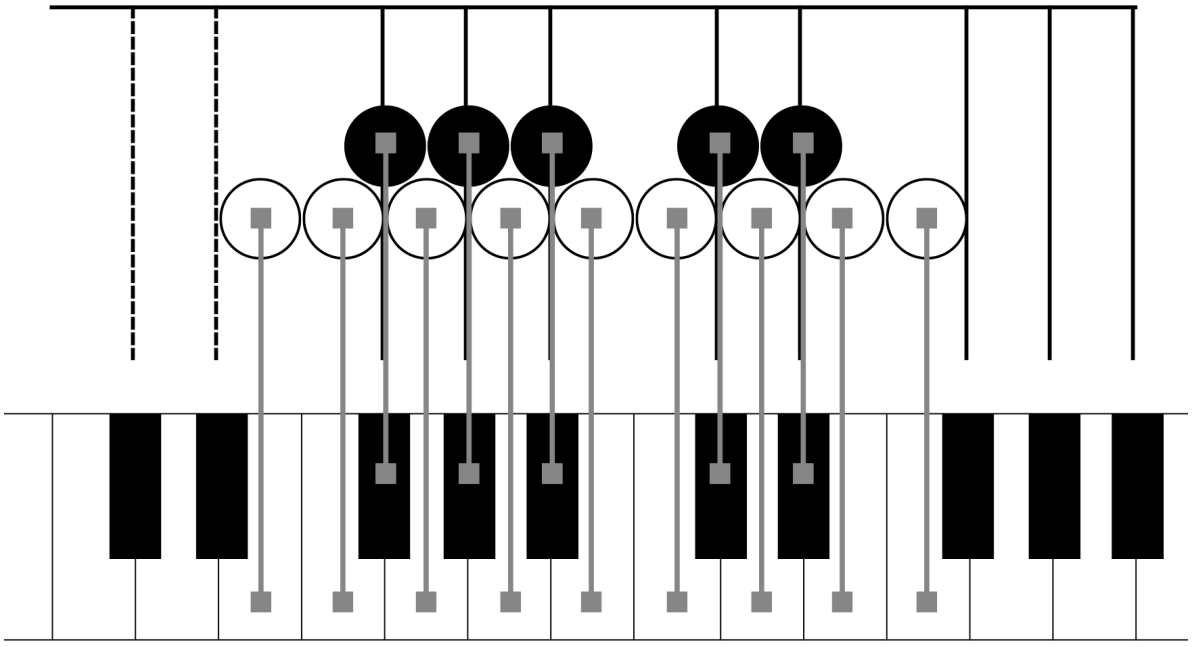
> *Note: Don't worry about the horizontal black line at the top of the staff. It is part of this notation system, but you may ignore it. You also can treat the dotted black lines on the left the same as the other vertical black lines; you do not need to understand the difference for the purposes of this guide.*

Every line and space between lines on the staff corresponds to a different key on the piano. Some spaces between lines are larger than others; these span two keys on the piano. The following diagram illustrates this correspondence.

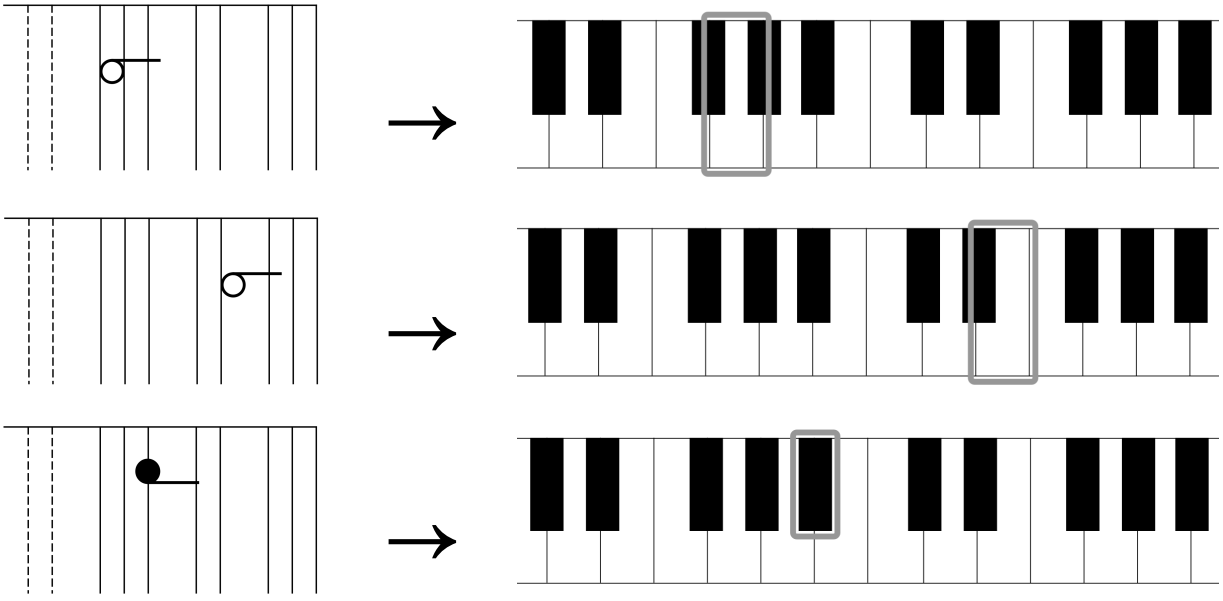


As shown above, the vertical lines of the staff correspond to the black keys of the piano, and the spaces between the vertical lines correspond to the white keys of the piano. The vertical lines come in alternating groups of two and three, just like the black keys of the piano.

A musical note, represented by a circle, is placed on one of the lines, or in the space between two lines. This indicates that you should play that note on the piano. The note is white if the corresponding piano key is white, and black if the corresponding piano key is black. Black notes are vertically offset from white notes to prevent notes from overlapping.
The following diagram shows several notes on the staff, and the corresponding keys on the piano.

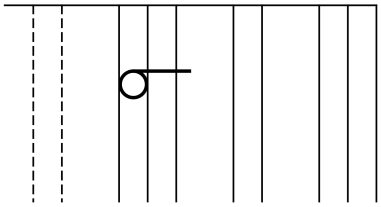


You will be asked to read a note on the staff, and identify the corresponding key on the piano. Below are several examples.



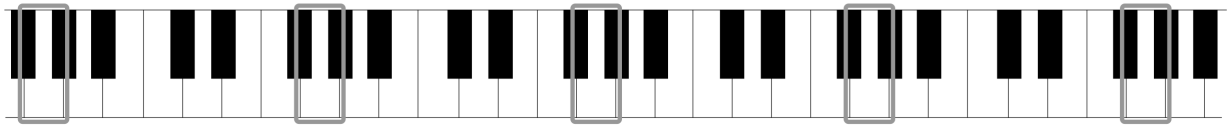
Note: Don't worry about the black lines extending to the right from the notes. These are part of the notation system, but you may ignore them.

For the purposes of this guide, when you are asked to play notes, do not worry about *where* on the piano you play these notes, as long as you play the correct keys within the pattern of black and white keys. For example, if you read



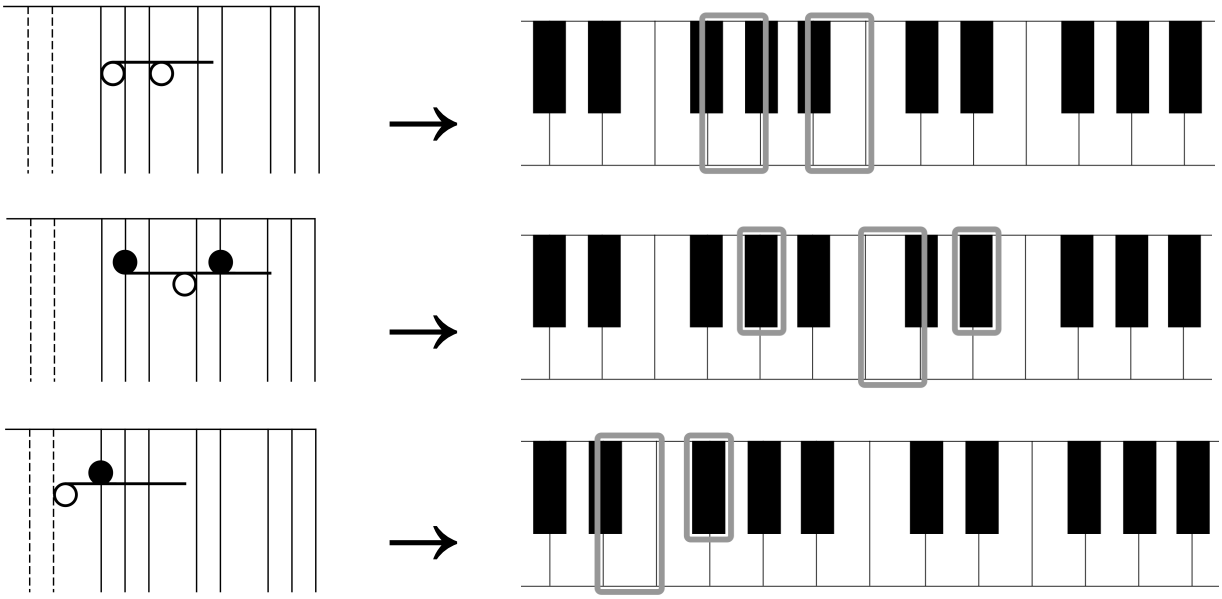
,

you can play any of the following keys:



Playing chords

A chord is two or more notes, played at the same time. Notes in chords are represented the same way as single notes. The notes are simply shown next to each other.



\- - -

This concludes the introduction to this music notation system.