VISUAL CONFUSION IN PIANO NOTATION

Marion Wood
University of Muenster
mwood@uni-muenster.de

ABSTRACT

This series of Reaction Time experiments investigates how quickly notes can be read from a screen and immediately executed on a MIDI keyboard. This makes it possible to study pitch reading and motor coordination in considerable detail away from the customary confounds of rhythm reading or pulse entrainment. The first experiment found that reaction times were slower in extreme keys (3#, 4#, 3b, 4b), even for very experienced sight-readers, a large effect of clef in most individuals, and other results suggesting that, in this simple paradigm at least, reading notation presents more of a difficulty to execution than motor coordination. A second experiment found, in addition, an effect of order in which the notes were presented.

A clarified form of notation was devised that disambiguates visual confusion across key signatures, and to some extent across clefs. Initial results from an experiment to contrast traditional noteheads with the clearer ones found substantial improvements in both Reaction Time and accuracy for the clarified notation. The possible applications of improved notation to the wider field of piano playing are discussed.

1. INTRODUCTION

Existing research into piano sight-reading [1] suggests that expert sight-readers may process common musical configurations as ‘chunks’ to a greater extent than novices. This study looks at the question of how some common musical chunks are learned or recognised.

Musical notation, considered as a semiotic system, is not a very effective map of the physical space of the piano keyboard (Figure 1). It does not illustrate the octave repeating pattern of the keyboard, and identical visual symbols or clusters of symbols must be executed differently by the two different clefs/hands.

Simply tabulating the different possible responses to a single common triad, (Figure 1) we find no less than ten visual-to-spatial mappings, considered across two clefs and eleven key signatures. The mappings also have different musical meanings: major, minor and diminished are words describing the musical ‘character’ of a chord.

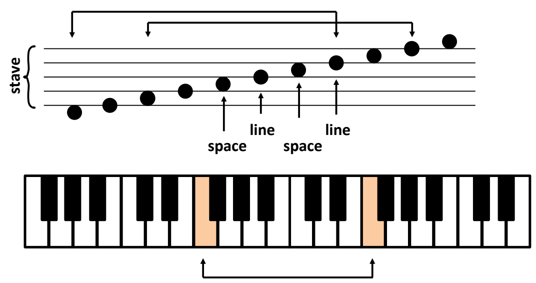

Figure 1. After seven notes of the scale, the keyboard repeats. Unfortunately the binary structure of the stave does not represent the number seven very effectively.

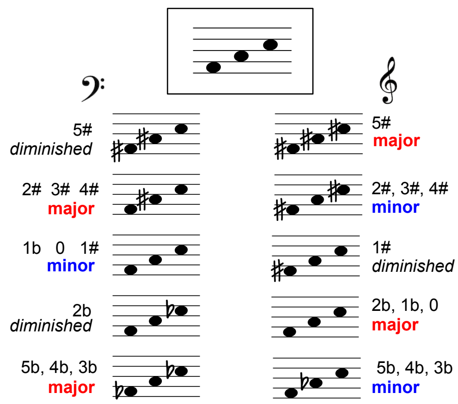

Figure 2. Ten different musical ‘meanings’, each with a specific motor response pattern, represented by a single visual fragment.

They all sound different, despite looking the same. The notation is not supporting ‘chunk’ learning or recognition by reflecting either execution mapping, or auditory mapping, or musical meanings.

Drawing a parallel from text reading, homographs and homophones cause particular difficulty for dyslexic readers [2]. Homographs are words that look the same, but whose sound and meaning are different: *lead*: *to go first* or *lead*: *a metal*. Conversely, homophones are words that sound the same, but whose meanings and visual presentations are different: *two, to, too*.

Even in non-dyslexic adult populations, homographs are read more slowly than singular control words, although homophones may be read marginally quicker [3]. Fortunately in most languages, these awkward words are the minority exceptions. By contrast, in piano music, any potentially recognisable musical ‘word’ – a chord, a scale fragment, or melodic pattern – can be classified both as a homophone and a homograph, having two separate execution patterns within any given key (Figure 3).

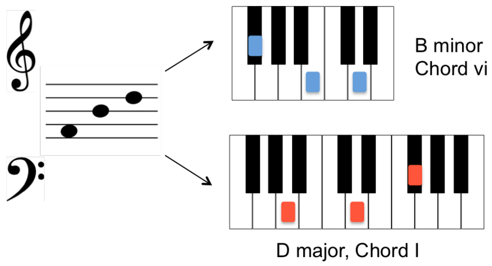

Figure 3a. The same visual fragment requires a different motor response in treble/bass clefs (right/left hands). Example from Key signatures of 1#, 2# or 3#.

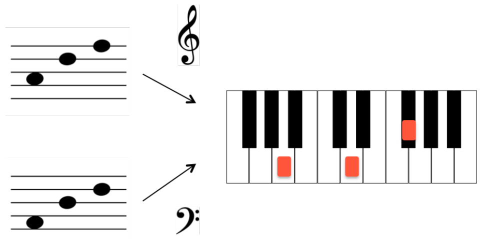

Figure 3b. A similar execution configuration and musical meaning requires two different visual presentations in treble/bass clefs (right/left hands). Example from Key signatures 1#, 2#, or 3#.

In mental chronometry research, visual processing is a topic of interest. Participants might be required to classify a visual stimulus according to various different rules, pressing one of two (or more) buttons in response, as quickly as possible. The time from stimulus presentation until the participant responds is the Reaction Time (RT).

Findings from this area include that are relevant to a discussion of sight-reading include: an increase in RT if the rules for responding are changed (a task-switch cost), longer RTs if the stimulus can be interpreted under two different rule-sets, a general increase in RTs when more than one rule set has to be held in mind at any one time [4], and the 'Simon effect' - an increase in RT if the buttons are arranged in an incongruent way, such as being required to press a right-hand button when a leftwards arrow is presented [5].

All of these factors may be considered to apply to music reading at the keyboard, where a left-right mapping on the keyboard is represented by low-to-high visual (and sound) mappings, and focusing on two different clefs requires us not only to hold two rule sets in mind, but also to switch between them frequently. The experiments below use standard RT paradigms to investigate these effects directly.

Some common musical patterns are normally taught to students of the piano under the topic of "Scales & Arpeggios", [6] although these exercises are often memorised. Thus while we may find that these patterns have been systematically rehearsed in their motor-execution, perhaps their recognition from visual presentation has not. Nevertheless, they represent the kind of chunks that we would expect expert sight-readers to recognise easily.

In the key signature 2#, for example, the chords of D major and B minor described in Figure 2 are so common in the musical literature that we would expect these patterns to become familiar very quickly to anyone who had played one or two tonal pieces in that key. Rather than asking how excellent sight-readers learn their skill, we should perhaps be asking why it is that so many pianists with years of experience do not. The hypothesis of this study is that overlapping visual representations may be part of the reason.

In summary, the experiments described below were designed to measure the reaction times of amateur and professional participants to visual musical stimuli in several keys and both clefs. Variations in reaction time were expected to reflect difficulties of processing the visual information, and/or motor coordination.

2. EXPERIMENT ONE

2.1 Method

Participants were requested to respond on a MIDI keyboard by playing series of 3-note combinations shown on a computer screen. Treble clef / right hand stimuli were shown in the top half of the screen, and Bass clef / left hand in the lower half, as seen in Figure 5. Participants were requested to play the notes in the order shown, as quickly as they were able to. There was no aural feedback (the keyboard was silent) but any errors were marked with red crosses on the screen after each trial.

The initial version of this experiment used a classic alternating task-switching paradigm [5], where two trials in one clef were followed by two trials in the other clef, in blocks of approximately 40 trials. The key signature remained the same for two blocks in a row and then was changed, with the whole experiment covering nine keys. (624 trials per participant). All stimuli were common triads in root position or inversion, all ascending, as shown in Figure 4.

The stimuli were grouped into three sets, with each block containing a mixture of two of these sets, while the other set rested. The order of presentation always followed a predictable pattern, *e.g.* (1)treble, (2)treble, (3)bass, (4)bass. Positions (1) and (3) are considered 'task-switch' positions where the clef has just been changed, and (2) and (4) 'task-repeat' conditions. Each stimulus was presented at some time during the block exactly once in each of these four positions. The order of stimulus presentations was otherwise randomised.

22 participants were recruited for the initial experiment by word of mouth from a variety of musical communities, in and around Exeter in Devon, UK.

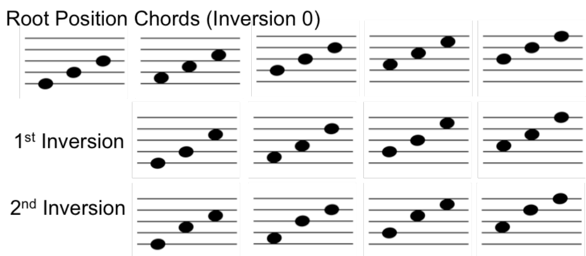

Figure 4. The 13 ascending triads that fall within the stave, used as stimuli in the initial experiment.

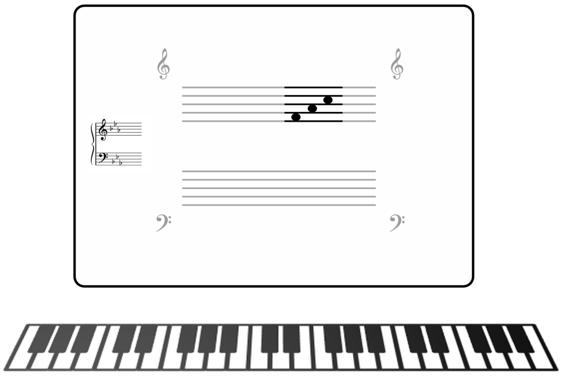

Figure 5. Experimental setup with screenshot of a repeat trial in the treble clef. A reminder key signature remains at the left during the whole block.

3. RESULTS AND DISCUSSION, EXPERIMENT ONE

3.1 Data Analysis

All 22 participants had a level of amateur involvement in music, and about half also had a professional component to their musical activities. They ranged in age from 18 to 74, and use the piano in a variety of different situations, including solo performance, teaching piano, teaching classroom music, teaching another instrument, accompanying another instrument, or learning music they later plan to sing.

The data analysis relies on averaging the mean RT over groups of participants across the cells of the design. Although it would be possible to normalise the data across all participants, there are some aspects of motor coordination and cognitive architecture which are common to all levels of competence. Reaction time is a direct reflection of a physical quantity (processing duration) and is consequently not usually transformed in reporting experiments of this type. Consequently the participants were divided into two main equal-sized groups, consisting of those with an average RT in the region 800-1500ms (11 participants) and those in the region 1500-2500ms (11 participants) with their data averaged in contrasting conditions of interest. As it turned out, this division into two groups on the basis of RT mapped onto a difference in musical history between those who have or had some professional component to their music and those for whom the piano was an adjunct to their other musical activities or a less serious hobby.

Various combinations of factors were grouped for analysis in repeated measures ANOVAs. Error scores of -1 were mostly single errors of execution in the correct hand in the right general area of the keyboard, whereas errors of type -3 were almost all mistakes of switching (the wrong hand used, or wrong clef read).

Main findings of the effect of clef, switch of clef, key signature, change of key signature and effect of the preceding presentation condition of a visual stimulus are reported in detail and discussed below. Other findings of the effect of inversion, diatonic chord and difficulty of hand execution are summarised more briefly.

3.2 Effect of Clef

3.2.1 Results

This contrast compared the mean reaction times found in the two clefs. In this experiment the treble clef was always played by the right hand and the bass clef by the left hand, and so any disparity might be caused either by differences in reading the clefs, or motor coordination differences between the hands, or a mixture of both. Across the expert group, the mean reaction times were treble/right, 1115ms, and bass/left, 1233ms: a difference of 118ms, $F(1,10)=31.77$, $p<0.001$. In the moderate group, these values were 1835ms and 2044ms respectively: a difference of 209ms, $F(1,10)=34.73$, $p<0.001$. Percentage errors were also greater in the bass clef for both groups, but this difference was not reliable in either group, either for the total or for any value of score.

Of 25 participants, three reported being left handed, and three would read music more often in bass clef outside of their piano playing, for example when playing the 'cello. One participant was in both of these groups. All these participants, however, performed significantly better in right hand/treble clef trials. In fact no participants were found for whom the left hand/bass clef showed an advantage compared to the right/treble.

3.2.2 Discussion

Left-handed participants expressed little surprise on being informed that their treble/right hand RTs were faster than their bass/left. They mostly reported the view that they had learned the treble clef first, and therefore had always felt more fluent reading it. In terms of accumulated reading practice, it is also the case that piano music tends to contain more notes in the right hand than the left. Consequently here is probably not the place for a wide-ranging discussion of handedness. However, this finding lends general support to the idea that reading the notation may be more of an issue than motor coordination.

3.3 Effect of Switch of Clef

3.3.1 Results

This comparison contrasted trials where the clef had just been 'switched' with those where the clef was repeated. Across the whole data set, a time cost of switching clef, as opposed to repeating the previous clef, was found. In the expert group the mean RT on clef switch trials was 1232ms, and 1186ms on repeat trials, a difference of 46ms, $F(1,10)=23.64$, $p=0.001$. In the moderate group these values were 2055ms and 1917ms; a difference of 138ms, $F(1,10)=13.75$, $p=0.004$. Levels of error were not significant in either group.

3.3.2 Discussion

During the course of the experiment, it became clear both by observation and self-report that a number of participants were finding it very difficult to maintain the pattern of two-trials-per-hand. Several experienced pianists appeared to be so thoroughly accustomed to alternating hands that they found it extremely hard, even after 20 minutes, to remember to repeat each clef. Eye-tracking studies of fluent sight-readers report a frequent alternation of saccades between clefs [7], [8]; this habitual pattern may be harder to shake off than expected.

Notwithstanding the unexpected difficulty in maintaining the predictable pattern of the experiment, a clear cost in reaction time of switching clef was found. Although not large in comparison to other effects found in this study, this result is interesting in the context of task-switching literature. After hundreds of hours training in task-switching labs, the question of whether participants can ever eliminate the switch-cost with sufficient practice is still hotly debated. This results suggests that switch-costs remain an issue in piano playing, even after thousands of hours of practice.

3.4 Effect of Key Signature

3.4.1 Results

Key signature as a whole was found to be highly significant. Participants generally performed more slowly in "extreme" keys, and faster in "central" keys: 1b, 0, 1#, as seen in Figure 6. In individuals, the pattern was influenced to a greater or lesser extent by favourite keys or recent experience, but the sensitivity to key signature was by far the most substantial effect seen in the experiment. Of all 22, 16 performed best in the key 0, and all but one of the others in either 1# or 1b.

In the expert group, performance was slowest in the key of 2#, with a mean RT of 1196ms, and fastest in the central key of 0, at 1011ms; a difference of 185ms ($F(8,80)=9.54$, $p<0.001$). In the moderate group the slowest average performance was in the key signature of 4#, (2298ms), and fastest in the key 0, (1678ms); a difference of 620ms ($F(8,80)=16.13$, $p=0.002$).

Individual preference or experience in key performance tended to cancel one another out in the means quoted above. In fact no participant's individual variation between their best and worst key signatures was less than 200ms. In the expert group, the mean of individual differences between best and worst performance in a key signature was 359ms, with the individual differences lying between 200ms and 750ms. In the moderate group, the average individual difference between best and worst keys was 795ms, with a range individual differences varying from 285ms to 1205ms.

Error scores showed no statistically reliable effects, but single-note errors showed some sign of approaching significance, and mirrored the shape of the key change variable: for the expert group $F(8,80)=2.04$, $p=0.108$, and for the moderate group $F(1,10)=3.94$, $p=0.075$.

3.4.2 Discussion

This is a very substantial finding: although sensitivity to key signature varies greatly, apparently even the most proficient pianists are not immune to its effects. The most experienced professional in the experimental set, with thousands of hours experience in playing, sight-reading and accompanying, had a mean RT on correct trials in the central key of 831ms, rising to 1064ms in 4# and 1062ms in 4b: a difference of 233ms. Expressed as a percentage of best performance, the effect of key signature appeared to add some 25% to reaction time.

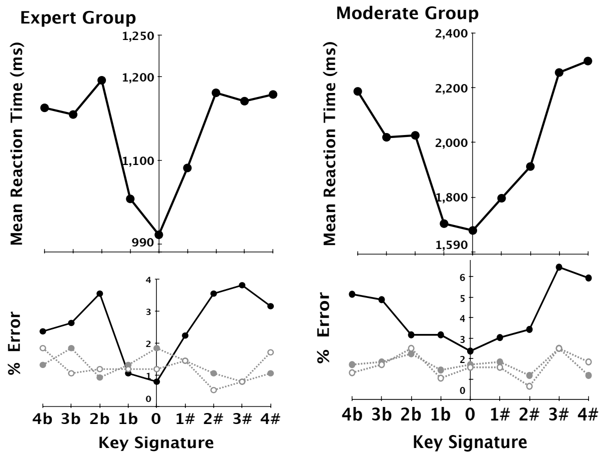

Figure 6. Average RT for Expert and Moderate groups across 9 common key signatures.

This is a result that would be surprising to most musicians, although perhaps not to researchers familiar with the mental chronometry literature. Pianists are generally supposed to become fluent "in all keys" with sufficient practice. The idea that an experienced professional might be as much as 25% slower processing pitch patterns in outer keys than in central ones runs counter to the prevailing view of practical proficiency in piano sight-reading.

In terms of a more nuanced pattern of key signature difficulty, it was seen that participants did not necessarily all find the outermost keys the most challenging. Indeed by self-report, when there were more than 3 modifiers in the key signature (4#, 4b), some participants used a strategy of remembering which black notes not to play (there are 5), rather than keeping track of all the modifiers. This resulted in some participants expressing the idea that outer keys of 4#, 4b, were actually easier than the "middle" keys of 3#, 3b. Participants of this type were more frequent in the expert group, which may be seen from the shape of the means plotted in Figure 6.

3.5 Effect of Changing Key Signature

3.5.1 Results

The key signature was changed every other block, and so reaction times could be contrasted between key change blocks, and those where the key remained as previously. Significantly higher average reaction times were found both the expert and moderate groups in the key change blocks, and so a finer grained analysis, dividing the blocks into thirds, was conducted.

A clear pattern of "settling into" the key signature was seen (Figure 7). In the expert group this was largely captured by a drop of 74ms in mean RT from 1308ms in the first third of a key change block to 1234ms in the next. In the moderate group the drop was 148ms, comparing 2205ms in first thirds of key change blocks to 2057ms in second thirds. The interaction of these effects was statistically reliable: in the expert group F(2,20)=8.45, p=0.002, and the moderate group F(2,20)=9.18, p=0.001. An analysis of errors did not reach significance.

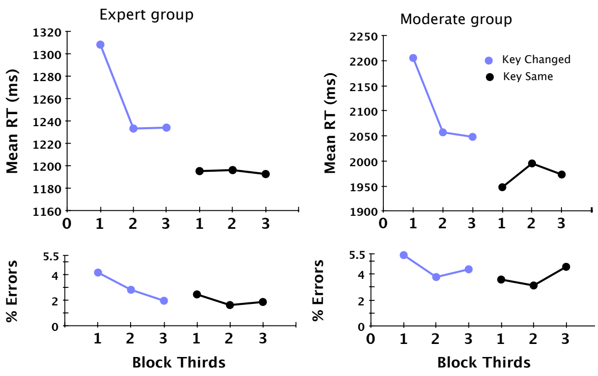

Figure 7. Means for each third of a block, for pairs of blocks in the same key.

3.6 Effect of Novel/Repeat Stimulus groups, and Clef-congruence of Previous Presentation

The effect of stimuli on one another within the experiment was analysed in two ways.

On a global scale, three subsets of stimuli were rotated so that half the trials in each block were from a 'repeat set' – i.e. they were also shown in the previous block, and half from a "novel set" that had been absent in the previous block.

At the local level, within each block, each stimulus appeared four times, once in each clef-switch/repeat condition, i.e. twice in each clef in each block. Investigating whether the RT of a stimulus is affected by its most recent previous appearance, trials were coded according to whether the stimulus had most recently been seen in the same (similar) clef, or in the other clef (different): see Figure 8. Stimuli most recently seen in a previous key signature were removed from this analysis.

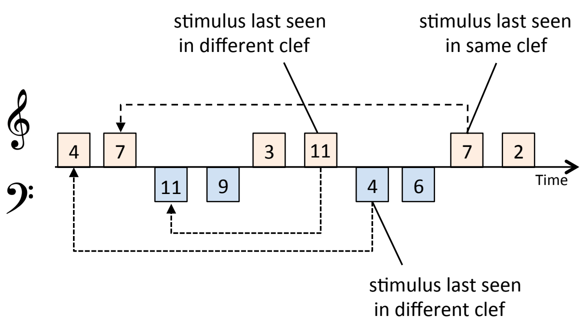

Figure 8. Illustration of last-seen-clef similarity. (Stimuli numbered arbitrarily).

3.6.1 Results

Comparing the two subsets of 'novel' and 'repeat' stimuli within blocks where the key signature remained the same, no significant effect was found in either the expert or moderate groups, or in the error rates. The variable describing 'last-seen-clef' congruence, however, was found to be highly significant in both expert and moderate groups. In the expert group, the mean RT of congruent last-seen-clef trials was 1176ms, whilst mean RT where the last-seen-clef had been different was 1220ms; a contrast of 45ms, F(2,20)=27.91, p<0.001. In the moderate group, the mean RT for congruence of last-seen-clef was 1910ms, and for incongruent last-seen-clef 2006ms; a contrast of 96ms, F(2,20)=13.25, p=0.001.

3.6.2 Discussion

This is an important finding. Practicing one visual stimulus (albeit with two different hand interpretations) might be expected to have an effect on the same stimulus in further blocks of the same key. Having either learned, or been 'reminded' of how a particular visual sign should be executed in both hands, we might reasonably expect an improvement in performance in the second block of the same key signature. The fact that any such improvement was not detectable in this experiment, whilst instead, the clef similarity of the most recent previous presentation did make a significant difference, suggests that visual confusion at the local level is active in a substantial way, and may be disrupting more general learning of patterns across both hands.

3.7 Other Effects

Other results are summarised in brief.

3.7.1 Effect of black notes

The number of black notes present in each chord could provide a simple reason for slower performance in outer keys, being perhaps harder to read or execute. Comparing triads with 0, 1, or 2 black notes required two contrasts, as not all types occur in every key. However no effect was found for the number of accidentals, except for a small but significant difference in the expert group, between chords with one and two accidentals, of 41ms, seen in Figure 9.

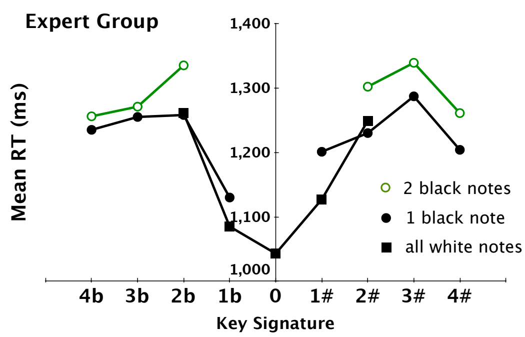

Figure 9. Mean RT in the Expert group of chords with 0, 1 and 2 accidentals.

3.7.2 Inversion

The effect of chord inversion (see Figure 4) was also analysed in combination with clef and the number of black notes in the chord. Unexpectedly, inversion turned out to be a significant factor in itself, and no reliable interaction was observed with hand (treble/bass clef) or number of black notes in the chord. For the expert group the mean RT between fastest (root position) and slowest (second inversion) was 29ms, and for the moderate group there was a much larger difference of 216ms. This effect was entirely unexpected, as inversion is relatively well-represented in the notation, with a slightly wider vertical gap in some triads (Figure 3) corresponding to a greater distance on the keyboard, and execution configurations not noticeably more complex.

Possible reasons for the effect of inversion include a bias towards recognising the chord by its root note, which in this experiment was the initial note presented in the case of the root position chord. Alternatively, it may be easier from a visual processing perspective to read three similar notes all on lines or all in spaces, than to distinguish a mixture of the two types.

3.7.3 Diatonic Chord

The seven chords of each key can be classified by musical type, commonly referenced in music theory by their roman numerals. In each key there are three major chords (I, IV, V), three minor chords (ii, iii, iv) and a single diminished chord (*vii*). Of these, I and vi are the 'naming' chords of each key. Means for these diatonic chord types are shown in Figure 10.

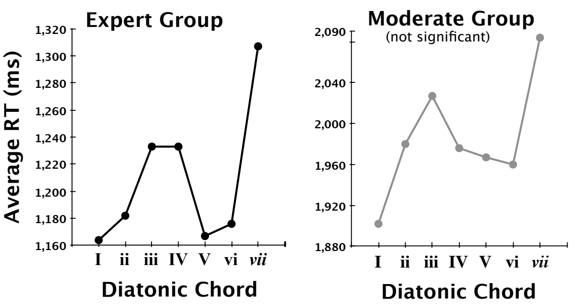

Figure 10. Variation in mean RTs according to diatonic chord. I is the major key chord, vi the minor key chord.

The difference between chord *vii* and the mean RT of all the other chords in the key was significant in both groups, with mean RT differences of 115s and 116s in the expert and moderate groups respectively. This chord is less common in the literature, and requires a slightly different hand configuration.

In the expert group pairwise comparisons between individual diatonic chords were significant for chords I, iii, IV and vi. The idea that chords I and vi may be more easily recognised, particularly by expert participants, is broadly encouraged by these results.

3.7.4 Single-Note Errors

Responses in which only one note was played incorrectly were collated, and showed that across the whole experiment, about 50% were caused by omitting the last accidental of the key signature – the 4th note of the scale in flat keys and the 7th note of the scale in sharp keys. This effect seemed to be irrespective the order in which key signatures were presented.

3.7.5 Unbalancing of the Design

The effects of Inversion and Diatonic Chord unbalanced the design of this experiment to some extent. In the total set of 13 stimuli, there were 5 root position chords, and only 4 of each of first and second inversions (Figure 3). The stimuli were drawn randomly to form subgroups that would rotate across blocks, and there was no control to ensure that an approximately equal mix of inversion types or diatonic chords fell in each block, or consequently each key signature. The larger effects of clef and key signature were likely to be valid findings, but balanced sets should be a requirement of any future experimental design.

3.8 Summary

Table 1 shows a summary of results from experiment one. As noted in the introduction, existing literature suggests that pattern recognition may play a part in sight-reading fluency, and this is broadly supported by these results. The effect of inversion appears to reduce substantially with expertise, and the differentiation between diatonic chords to increase. That clef and key signature should remain such large sources of variation in the expert group is at odds with the prevailing impression amongst musicians. The only good evidence for motor-coordination challenges is provided by the expert group in executing chords with two black notes. Other findings generally support the idea that decoding multivalent notation may be a substantial challenge faced by all pianists, more-or-less regardless of expertise.

Effect	**Expert**	**Moderate**
range of average RTs	*800-1500ms*	*1500-2500ms*
Clef	118 ms	209
Switch of clef	46	138
Key signature	185	620*
Effect of key change	65	111*
Last-seen-clef similarity	45	96
2 black notes	41	(n.s.)*
Inversion	29	216
Diminished chord	115	116*
Other diatonic chords	70*	(n.s.)*

* results that may be unbalanced by the effect of inversion/diatonic chord

Table 1. Summary of results from experiment 1, with differences in mean RTs given in ms.

4. EXPERIMENT TWO

4.1 Method

A follow-up experiment took place at Dartington Summer School, likewise recruiting interested volunteers by word-of-mouth, at a variety of expertise levels. Participants were sorted according to performance on a practice block, with 15 moderate participants completing a shorter, balanced version of the experiment with 252 trials.

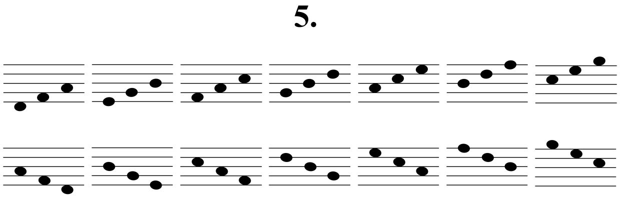

Figure 11. 14 musically balanced triads used as stimuli in a second experiment.

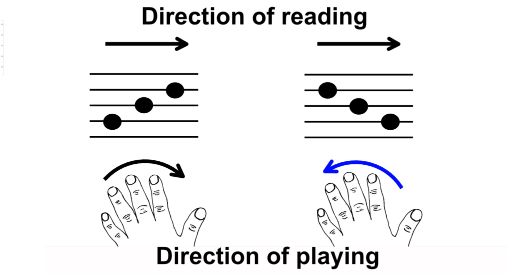

Figure 12. Congruence and incongruence of playing direction *vs* reading direction, in forward and reverse presentation of stimuli.

This experiment omitted the task-switching element of the design, (which, although statistically significant, was not very large), and instead presented trials alternately to each hand.

Seven root position chords were used (see Figure 11), and every stimulus was presented in every key. In a variation of the original experiment, each triad was presented both forward, and in reverse, *i.e.* with the highest note first. The hypothesis was that incongruence of direction (Figure 12) might provoke a 'Simon effect' [5], with descending Figures disadvantaged.

6. RESULTS & DISCUSSION, EXPERIMENT TWO

6.1 Data Analysis

15 moderate participants completed the experiment, with RTs in the range 1150-3300ms. 12 of these fell in the range 1400 – 2400ms. Given that the experiment was restricted to root position triads, this group forms a good comparison with the moderate group of Experiment 1.

6.2 Effect of Clef

6.2.1 Results

The mean reaction times for contrasting clefs were treble/right, 1722ms, and bass/left, 1726ms: a difference of 109ms, $F(1,11)=15.69$, $p=0.002$. Once again, three self-reported left-handers showed no left hand advantage.

6.3 Effect of Key Signature

6.3.1 Results

As before, key signature was found to be significant. The slowest average performance was in the key signature of 4b, (1973ms), and fastest in the key 0, (1488ms); a difference of 485ms ($F(8,99)=6.10$, $p<0.001$).

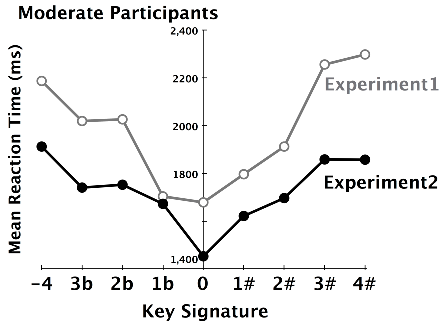

Figure 13. Mean RTs across 9 common key signatures; data from both experiments for comparison.

6.3.2 Discussion

Although, once again, individual key signature profiles showed great diversity, the averaged data for this balanced version of the experiment showed interesting similarities to data from the previous experiment, seen in Figure 13. The 'kink' in the flat keys suggesting an advantage for 3b, and what may be a corresponding disadvantage for 3# are apparently consistent features that would bear further investigation.

6.4 Effect of Order

6.4.1 Results

The order of presentation of the triads was statistically significant, with a mean reaction time of 1726ms in ascending triads, and 1830ms descending, a difference of 104ms $F(1,11)=36.71$, $p<0.001$. No interaction with clef was found. Across the whole experiment, single-note errors showed a tendency to be more common in the third note than the first two: of 84 such errors, 45 were in the third note.

6.4.2 Discussion

This result is similar in size to the effect of clef, and may to some extent reflect the cognitive architecture required to process a mirror rotation. While not discounting the possible hypothesis that presenting the naming note of the chord first confers an advantage, the pattern of errors (either in this experiment or the previous) did not show any evidence that the key note was preferentially recognised.

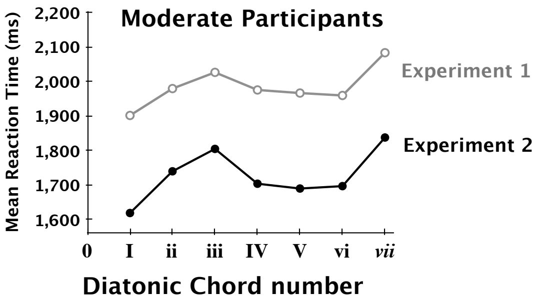

Figure 14. Mean RTs across 7 diatonic chords.

6.5 Effect of Chord Type.

In this experiment an analysis of diatonic chord was significant, with lowest mean reaction time for chord I 1662ms, and the highest for the diminished chord *vii* of 1895ms, a difference of 233ms, $F(6,33) = 8.57$, $p<0.001$. The graph of mean RTs for individual diatonic chords was again very similar to the corresponding previous data; see Figure 14.

6.6 Summary

The follow-up experiment clarified previous results of key signature and diatonic chord using a more rigorously balanced design, and found an effect of order of notes to be played. The particular shapes of the key signature and diatonic graphs are interesting and merit further exploration.

7. EXPERIMENT THREE: IMPROVED NOTATION SYSTEM

There have been various attempts to improve piano notation to be better suited to describing the execution of music at the keyboard[1]. There is great resistance to any kind of notation, however, that does not take account of the enormous canon of existing literature, or the years of investment by current professionals in the traditional system. Is any modification to 'standard' notation possible that might clarify the cognitive difficulties, whilst remaining legible to those accustomed to traditional notation? One suggestion is given in Figure 15.

Figure 15. Modified noteheads, showing the chords Bb major and D major. Notes with the left half filled are flat (b), and with the right half filled, sharp (#). Short barlines are also placed to clarify the clefs.

Modification of noteheads had been implemented in the testing program from the outset, partly to provide a fallback for those who found the main experiment beyond their capability. Assorted pilot data indicated that this notation improved RTs at every level of competence, and in the case of one or two dyslexic participants made a transformational difference. The improvement was difficult to quantify simply by contrasting complete runs of the experiment, as there was also a considerable learning effect whenever the experiment was taken more than once.

A third experiment recruited participants mainly from the German town of Münster, mostly in the age-group 18-30, from the University Choir or Institute for Music Education.

Blocks of the clarified notation shown in Figure 15 were presented alternately with blocks of the black noteheads used in previous experiments. Six 'difficult' key signatures (4b,3b,2b, 2#,3#,4#) plus the central key of 0 were arranged in one of four maximally confusing key orders. Each key signature was presented twice in each run of the experiment, in such a way that all six key signatures were seen in both notations. In other respects the procedure was identical to Experiment two, apart from including one extra practice block of the new notation. The overall design investigates the enabling or disruptive effect of key signatures on one another.

8. RESULTS & DISCUSSION, EXPERIMENT THREE

8.1 Data Analysis

At the time of writing, 10 moderate participants with average RTs between 1400 and 2450 had completed two contrasting runs of the experiment, as one quarter of a larger 4x4 design. (Data from a further thee expert participants, plus eight who completed one experiment are not reported here.) Runs were undertaken in the same session, with not more than 20 minutes break between them.

8.2 Summary of Replicated Effects

The effects of clef, order and key signature were consistent with the previous experiments, with an average clef difference of 203ms, a difference between slowest and fastest keys (4#, 0) of 623ms, a difference between rising and falling note orders of 151ms, all highly significant, and a difference between diatonic chords (I and *vii*) of 247ms. The diatonic profile showed a relative advantage for chord V compared with previous results, making it slightly more like the expert profile seen in experiment one: see Figures 17, 14 and 10.

8.3 Effect of Clarified Notation

The effect of clarified notation was analysed in combination with the other factors across the six difficult keys.[2] The average RT for traditional notation (all black circles) was 1970ms compared to 1693ms for clarified notation (see Figure 12), a difference of 277ms. ($F(1,60) = 22.1$, $p=0.001$). There was also a dramatic effect of notation on error scores. Results are shown in Figure 16.

---

[1] Klavar Notation, for example, is a well-developed alternative.

[2] Clarified notation could also be expected to improve performance in the central key, by removing the conflicting mappings from other keys, but not in an experiment where traditional notation is also being presented.

There was also no interaction with the effect of diatonic chord, shown clearly in Figure 17. Both clarified and traditional notation showed similarly ‘musical’ patterns.

8.4 Learning Effect

8.4.1 Results

There was a significant learning effect across the two experiments. Across the difficult keys, the average RT for the first run was 1923ms, and 1724ms for the second; a difference of 199ms, $F(1,54)=43.28$, $p<0.001$. (This compares to the learning effect in the central key (0#/b) that fell from 1459ms to 1304ms; a proportionately comparable drop of 155ms.) The effect did not show an interaction with the clarified notation, which appeared to confer a similar advantage across the two experiments of 279ms and 274ms.

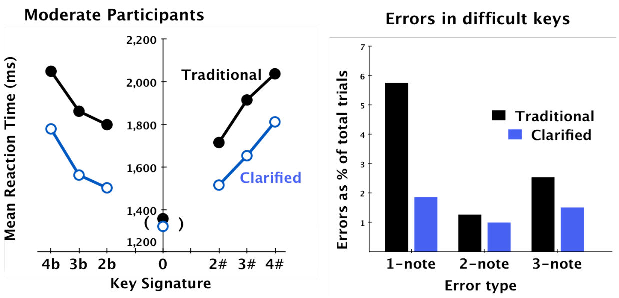

Figure 16. Results in difficult keys: comparing traditional and clarified notation.

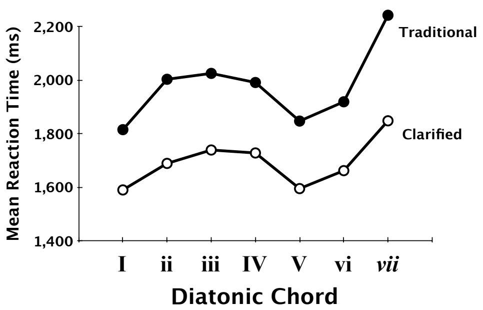

Figure 17. Results in difficult keys: Diatonic chords.

Rates of error also remained constant between the experiments, excepting 3-note-type errors, which fell slightly in the second experiment in both notations.

8.5 Participant feedback

8.5.1 Results

All participants commented that they found the clarified notation easier/faster to process, despite having been advised that the main aim of the altered notation in this experiment was to test the effect of key signatures on one another. Participants remarked that they not only performed more quickly, but were also more sure of their answers, and therefore felt less requirement to double check every response for errors before pressing the keys. A number of them made unprompted suggestions about how the clarifications could be introduced into ordinary piano music, notwithstanding the need to differentiate minims (whole notes) from crotchets (quarter notes).

8.6 Discussion

These are large effects in cognitive processing terms, with a gain of 15-20% on reaction times in difficult keys, and a halving of errors. More data is needed to complete the contrast of the particular key signatures being studied, but clearly from an experimental design perspective this is a useful way to separate some of the visual effects from other features of motor architecture or cognitive musical structure. It required very little acclimatisation, and conferred what appears so far to be a consistent advantage in difficult keys.

In terms of incorporating some clarifications into standard Piano music, the comments from participants are interesting. It may be argued that the visual disadvantage of overlapping pitch representations is somewhat overstated in these experiments, as there is so much extra contextual information on a real piano score. Looking at the question in reverse, however, freeing attention and working memory from the constant over-checking for pitch errors could leave room for sight-readers to take in more of that contextual information, resulting in bigger gains than those reported here.

It is encouraging that the diatonic key profile appears to persist in the clarified notation, as a further objection would be that clarifying the notation would reduce sight-reading to simple ‘button-pushing’ without the need for any musical understanding. In fact there was some indication from earlier experiments that at least part of this musical structure learning takes place outside of conscious theoretical understanding; some participants could not name either the major of minor keynote of most key signatures but nevertheless showed data of approximately this pattern. Clarifying the notation across the standard repertoire might simply have the effect of accelerating the pattern-learning process.

9. CONCLUSIONS AND DIRECTIONS FOR FURTHER STUDY

This study demonstrates that even sight-readers who excel have not achieved equal familiarity with every key, or parity of reading/execution between the clefs, and begins the process of investigating why this might be so.

The nature of the overlapping mappings, and the effect of inversion and diatonic chord make it complex to disentangle the effects of one clef on the other one, one key signature on another, or one visual pattern on another without finding a way to remove some of the confounds. Clarified notation creates a useful contrast to disentangle some of these effects and may itself provide either a training aid, or a structured alternative to traditional notation for those who find it useful.

9.1 Directions for Further Study

The third experiment uses clarified notation to investigate particular aspects of interference between one key and another, with data collection continuing at the time of writing. Further study is needed to discover whether the learning effect seen in traditional notation blocks is improved by interspersing blocks of clarified notation, or proceeds independently of it.

Further work also aims to involve the dyslexic / dyscalculic population, for whom Piano sight-reading often presents a disproportionate challenge, and for whom an alternative notation could offer particularly relevant benefits.

Presenting a form of clarified notation in a more realistic score format, and comparing attempts to sight-read simple pieces across more than one experimental session is also planned. Better sight-reading accuracy, and also better session-to-session retention in clarified notation is cautiously expected.

Acknowledgments

Many thanks to all the participants for their time, interest, suggestions and anecdotal insights, to Professor Stephen Monsell for his supervision in devising the experiment, and to the staff of Dartington Summer School for making the second experiment possible.

10. REFERENCES

[1] J. Sloboda, "The Eye-Hand Span-An Approach to the Study of Sight Reading," *Psychology of Music*, vol. 2, no. 2, pp. 4–10, Oct. 1974.

[2] Z. Breznitz, "Asynchrony of visual-orthographic and auditory-phonological word recognition processes: An underlying factor in dyslexia," *Reading and Writing*, vol. 15, no. 1–2, pp. 15–42, 2002.

[3] L. Gottlob, "Reading Homographs: Orthographic, Phonologic, and Semantic Dynamics," Journal of Experimental Psychology: Human Perception and Performance, Vol. 25, No.2, pp. 561–574, 1999.

[4] S. Monsell, "Task switching," *Trends in Cognitive Sciences*, vol. 7, no. 3, pp. 134–140, Mar. 2003.

[5] M. C. DeSoto, M. Fabiani, D. C. Geary, and G. Gratton, "When in Doubt, Do it Both Ways: Brain Evidence of the Simultaneous Activation of Conflicting Motor Responses in a Spatial Stroop Task," *Journal of Cognitive Neuroscience*, vol. 13, no. 4, pp. 523–536, May 2001.

[6] *Piano scales & arpeggios (from 2009). Grade 8. Grade 8.* London: ABRSM Pub., 2008.

[7] S. Furneaux and M. F. Land, "The Effects of Skill on the Eye-Hand Span during Musical Sight-Reading," *Proceedings: Biological Sciences*, vol. 266, no. 1436, pp. 2435–2440, Dec. 1999.

[8] Goolsby Thomas W., "Eye Movement in Music Reading: Effects of Reading Ability, Notational Complexity, and Encounters," *Music Perception: An Interdisciplinary Journal*, vol. 12, no. 1, pp. 77–96, Fall 1994.