Reading Pitches at the Piano: Disentangling the Difficulties.

Marion Wood

Institut für Psychologie, Westfälische Wilhelms-Universität, Münster, Germany
`mwood@uni-muenster.de`

ABSTRACT

To disentangle the cognitive processes involved in reading the pitch component of piano music, this research uses simple Reaction Time (RT) paradigms, and draws on task-switching and visual processing research. Participants were 44 pianists of widely varying age and competence. Pitches were shown three at a time and executed as fast as possible at a MIDI keyboard, with the key signature changed after 40-60 trials. A variety of specific factors could be quantified and contrasted, ranging from the cognitive 'switch-cost' of switching between clefs, to the extra processing time required when notes are presented in a different order. Large effects of clef and key signature persist in even the most accomplished professionals, adding 20% to response times at the very least, despite thousands of hours of practice. This paper tackles some of the questions about why this might be so. The extent to which overlapping, multivalent notation may obstruct both immediate execution and the longer-term learning process is discussed.

I. INTRODUCTION

Watching a pianist playing in a recital, often without any printed music, an audience member is often struck by the physical characteristics of the performance – the ability of the pianist to locate and play many notes in a complex and apparently effortless sequence. At the other end of the professional spectrum is the 'reluctant organist' attempting to play four hymns in a church service, with or without the added difficulty of coordinating notes on a pedalboard. We could be forgiven for thinking that motor coordination is the primary issue in studying a keyboard instrument. However a major issue preventing the exploration of a wide range of music that lies *within* the musician's technical grasp appears to be the difficulty of reading the printed notation. See Fourie, 2004, for example.

One of the problems in studying music reading is the interaction of the temporal (rhythmic) and physical (pitch) dimensions. This study takes as its focus the pitch aspect alone. By omitting the rhythm dimension, it is possible to apply insights from the considerable literature of existing response time (RT) research to the issue of music notation.

Musical notation, considered as a semiotic system, is not a very effective map of the physical space of the piano keyboard. 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.

Existing research into piano sight-reading (Sloboda, 1974) suggests that expert sight-readers may process common musical configurations as 'chunks' to a greater extent than novices. However as we see below, the notation system is not contributing effectively to the recognition of common musical patterns.

Tabulating the different possible responses to a single common triad, for example, (Figure 1) there are 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. They all sound different, despite looking the same. The notation is not reflecting differences in execution mapping, auditory mapping, or musical meaning.

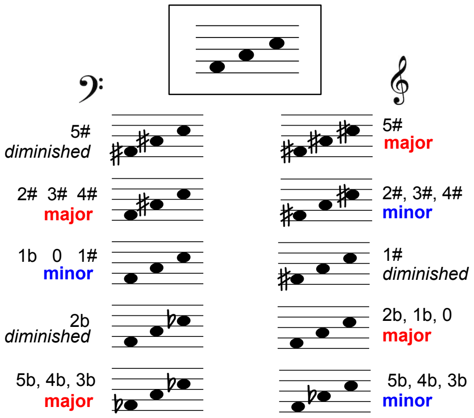

Figure 1. Ten different musical 'meanings', each with a specific motor response pattern, represented by a single visual fragment.

Research into the mental chronometry of visual processing suggests that 'multivalent' stimuli – *i.e.* visual symbols that can have more than one possible meaning – are processed more slowly than univalent symbols. Frequent changes to the response rules ('task-switches') slow the response time further, particularly in the first response after a switch. (Vandierendonck 2010).

Eye-tracking work by Goolsby (1994) showed pianists' gaze moving between treble and bass clef very frequently, sometimes on alternate saccades. If the two rule sets for treble and bass clefs can be considered a task-switch, clearly this would have implications for the speed of responding in piano sight-reading.

II. METHOD

Three separate experiments used a similar experimental setup, with three pitches presented to a computer screen at a time. The participant was requested to respond by playing the notes in the correct order, but as quickly as possible, as soon as they saw the stimulus. The key signature was given at the beginning of a block of 40-50 trials, and feedback on average response time and errors was given at the end of each block.

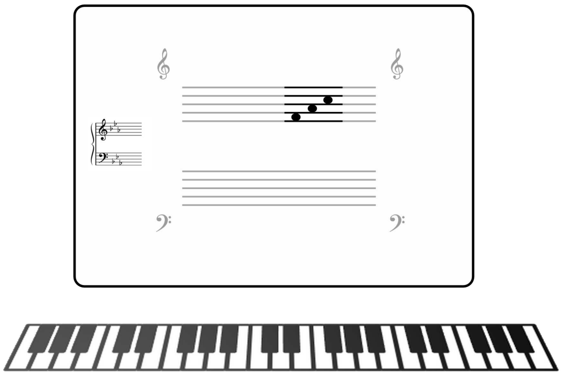

Figure 2. Experimental setup with screenshot of a repeat trial in the treble clef. A reminder key signature remains at the left during the whole block.

A. First Experiment

The first experiment used a classic alternating runs task-switching paradigm (Monsell 2003), in which two trials were presented to one hand/clef, followed by two trials to the other. Task-switch trials (immediately following a change of clef) may then be compared with task-repeat trials to measure the effect of the changing the task. A set of visually balanced triads was used as stimuli, shown in figure 3. This total set was subdivided so that not every triad appeared in every block, and the key signature was changed every two blocks. Thus a contrast could also be made between triads that had been recently rehearsed in the same key, or had recently appeared in a different key.

The whole experiment contained 18 blocks covering 9 keys, and were presented in two main key signature orders. 22 participants of widely varying age and musical background took part, recruited by word of mouth from in and around Exeter, UK, and their data were partitioned later into two equal-sized groups, of expert or moderate proficiency.

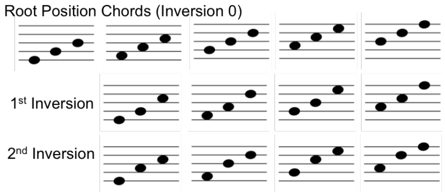

Figure 3. The stimulus set of experiment 1.

B. Second Experiment

The second experiment removed the task-switching element of the design (which although significant, was not found to be large in comparison to other effects) and also presented a more rigorously balanced set of experimental stimuli (figure 4), using seven root position chords. Every stimulus appeared in every block of every key. In this experiment each triad was also presented in reverse order. Nine key signatures were given in a randomly shuffled order.

Participants were all attending Dartington Summer School and were recruited through the summer school choir, so that in addition to their piano skills they were also all reasonably experienced choral singers. 12 participants had average RTs that approximately matched those of the moderate group from the previous experiment, and their data was analysed for comparison.

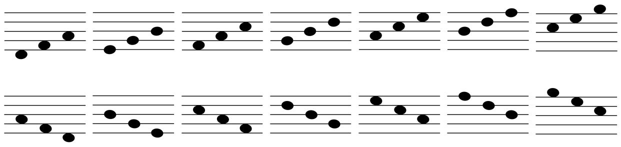

Figure 4. The balanced stimulus set of experiment 2.

C. Third Experiment

A third experiment focuses on the more difficult key signatures of 2#, 3#, 4# and 2b, 3b, 4b. Taking the same basic stimulus set as the second experiment, alternate blocks in the third experiment use a form of clarified notation (figure 5) that resolves the confusion between multivalent stimuli. 15 blocks present the six difficult key signatures (plus the central key C major/A minor) in one of four maximally confusing key orders, to try to evaluate the effect of one key signature on another, and establish to what extent visual ambiguity adds to response time.

Data from an initial group of 10 moderate participants is reported here, recruited from in and around Münster, Germany.

Figure 5. 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 clarify the clefs.

III. RESULTS

A more detailed account of these results may be found in the proceedings of the TENOR 2016 conference, where the the effects and applications of clarified notation are further discussed. Results are summarised here in brief, framing a discussion of the confounds in standard music-reading.

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.

Response time was taken at the third keypress, and averaged data across correct trials for analysis.

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).

A. Average Response Times

Participants in the ‘expert’ group of first experiment had average response times between 800 and 1500ms, and those in the ‘moderate’ group had response times between 1500 and 2500ms. Participants from the second experiment classed as ‘moderate’ had response times between 1150 and 2300ms, and from the third experiment between 1400 and 2450ms.

B. Effect of Clef

The effect of clef was highly significant across all experimental and pilot data, with the Treble clef generally faster than the bass clef. This Treble clef advantage included a number of self-reported left-handers, bass singers and 'cellists who might be expected to read bass clef more fluently.

In the expert group of the first experiment the contrast between the averages for each clef was 118ms, and in the moderate group it was 209ms. In the moderate group of the second experiment the impact of clef was proportionally lower but still highly significant, at 105ms. The contrast in the third experiment was 202ms, closely comparable to the moderate group of experiment one.

C. Effect of Switch of Clef

During the first experiment, in the expert group in particular, a number of very experienced individuals found it very hard not to alternate between the clefs/hands, and frequently hesitated or moved the wrong hand slightly in repeat trials.

Nevertheless a small but highly significant effect of 'task-switch' was found, in which the average response time when the clef had just been changed was greater than the average response time in trials where the clef remained the same. This difference was 46ms in the expert group, and 138ms in the moderate group.

D. Effect of Last-seen Clef Congruence

In the first experiment, the effect of stimuli on one another within the experiment was analysed in two ways.

1) 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.

2) 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 6.

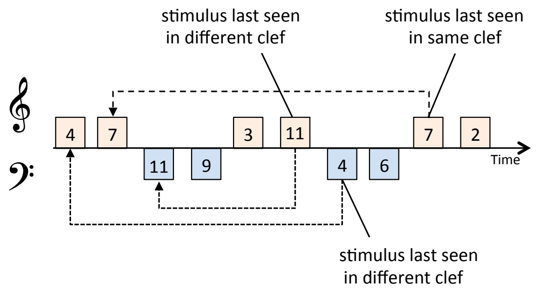

Figure 6. Illustration of last-seen-clef similarity. (Stimuli numbered arbitrarily).

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, with contrasts of 45ms and 96ms respectively.

E. Effect of Presentation Order

In the second and third experiments, the order of presentation was varied, with average response times to rising triads found to be faster than falling triads. In the second experiment this difference was 100ms, and in the third, 145ms.

F. Effect of Inversion

In the first experiment the effect of inversion was unexpectedly found to be significant, with no interaction with the number of black notes, or their position in the triad. In the expert group, this effect was small but highly significant, with a difference between the root and the first inversion 8ms, and between the first and second inversions 28ms, a combined difference of 29ms. In the moderate group, this effect was much larger, with a difference between root and first inversion of 168ms and between first and second inversions of 48ms, or 216ms combined.

Closer inspection of individual keypress data suggests that this may relate to the experimental procedure; the participants tended to arpeggiate their response in a single 'gesture'. The greater distance between two wider-spaced notes of an inverted chord is reflected in a slight delay between the keypresses. This effect appears to account for about half of the variation in inversion response times.

G. Effect of Diatonic Chord

The position of the chord within the key can be analysed according to its diatonic function, with chord 1 (I) the major key chord, and chord 6 (vi) the minor key chord. Chords 1-6 all appear in different positions in neighbouring keys, but chord 7 (*vii*), the diminished chord, is unique to each key signature and thus less practiced across the literature and within this experimental paradigm.

The contrast between chord 7 and the rest was significant in every group. The difference between the average response time for chord 1 and the average response time for chord 7 in the first experiment was 143ms in the expert group and 182ms in the moderate group; a proportionally much greater difference for the experts. In the second and third fully-balanced experiments this range was 219ms and 232ms respectively, with some pairwise comparisons between chords also significant.

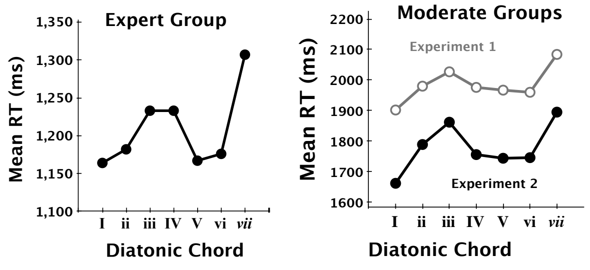

Figure 7. Diatonic Chord profiles

H. Effect of Key Signature

The contrast in performance between different key signatures was by far the largest effect seen in these experiments. There was a great variation in individual key signature profiles, but even in the expert group of the first experiment no participant showed less than 22% total variation across all nine keys, with others closer to 50%.

The first experiment was unbalanced by the unexpected effect of inversion, and a substantial effect of diatonic chord, neither of which had been allowed for in the selection of stimulus sets. In addition, the experts seemed to have continued to improve their performance beyond the practice blocks and into the first block of the experiment, which was either 2# or 2b. Consequently their key signature profiles, whilst generally indicative, may not be entirely typical. The second experiment was more rigorously balanced and shows a an approximately symmetrical shape, with key signatures gradually increasing in mean response time as the number of sharps or flats increases. Three modifiers seems to be an exception, with 3b showing what appears to be an advantage, and 3# a disadvantage. The third experiment, still in progress, sets out to investigate this apparent discrepancy further.

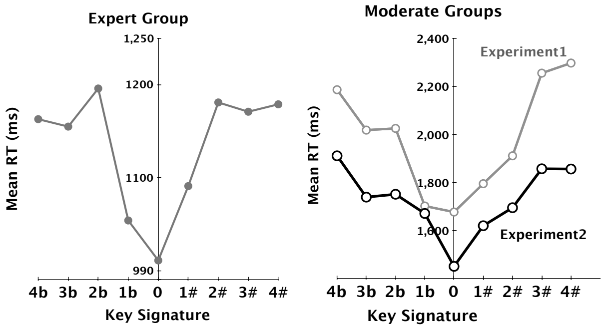

Figure 8. Key signature Profiles

I. Effect of Accidentals

An analysis of the number of black notes in each trial was used to test whether this factor could by itself account for the variation in key signature data. In the first experiment, in the expert group only, there was a moderately significant delay in executing chords with two black notes, but this effect was small in comparison to the general overall effect of keysignature. In the later experiments, although the same contrast showed signs of approaching significance, the effect was again very small, and in no sense accounts for the wide variation across key signatures.

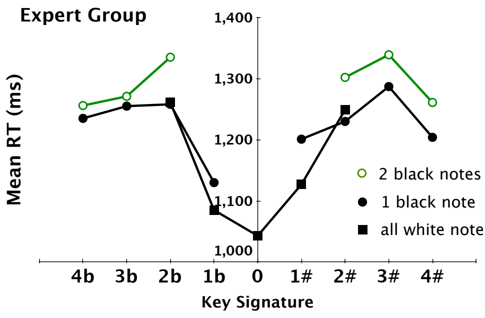

Figure 9. Mean RT in the Expert group of chords with 0, 1 and 2 accidentals.

J. Settling-in Effect

In the first experiment, in which two blocks of each key signature followed one another, the second block of each key was found to be faster than the first block. A finer-grained analysis found this settling-in effect took place mostly in the first third of a block: *i.e.* the first 11 or 12 trials.

K. Effect of Clarified Notation

The third experiment investigates the more difficult keys of 2, 3 and 4 modifiers and their effect on one another in close proximity. To elicit more information about the role of visual confusion in the notation, alternate blocks are given in the clarified notation of figure 3.

The effect of this notation was itself highly significant, with the average response time in clarified blocks 277ms faster than in the traditional blocks. In addition, a dramatic fall in the error rates was observed, particularly in 1-note errors. For more discussion, see the proceedings of the TENOR 2016 conference (in press).

For the current discussion, although the data presented here constitutes only one quarter of a larger 4x4 design, initial analysis suggests that 3# remains more difficult than 3b, and that neither the learning effect provided by an intensive balanced design, or the effect of clarified notation seems to have entirely dispelled this tendency. On the other hand, clarified notation seems to have mitigated about half the total effect of key signature.

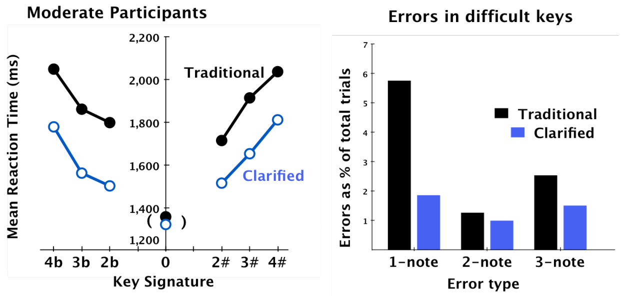

Figure 10, Effect of Clarified Notation

L. Summary of Results

Table 1. Summary of effects across three experiments. All times given in ms.

	Experiment 1		Exp. 2	Exp. 3
Effect	Expert	Mod.	Mod.	Mod.
Range of average RTs	*800-1500*	*1500-2500*	*1150-2300*	*1400-2450*
Clef	118	209	105	202
Switch of clef	46	138		
Key signature	185	620*	460	575
Effect of key change	65	111*		
Last-seen-clef	45	96		
2 black notes	41	(n.s.)*		
Inversion	29	216		
Diatonic Chord	143	182*	219	232
Key Signature	185	620*	460	575
Note Order			100	145
Notation				277

* Effects that may be unbalanced by the effect of inversion.

IV. DISCUSSION

In principle there may be a number of factors at work to explain the variation in response times reported here. Hand-eye coordination in this situation requires at least three steps: decoding of the instructions, some transformation into an action plan, and then the execution of that plan in space. Delays may be incurred by extra difficulties at any of these stages.

The only evidence for significant motor difficulty in executing common chords at the keyboard was provided by the very small delay for chords with two black notes in the expert group of the first experiment. The contrast between clef/hand is unlikely to be a coordination problem because it remained evident in left-handers.

A likely example of a transformation difficulty is the effect of presentation order (III.E). In the context of piano playing, ascending figures are congruent to the order in which the notes are to be executed, shown in figure 11 below. Processing notes from left to right and then executing them in the other direction requires a mirror transformation that may confer a delay.

The difference between the clefs themselves can also be considered from this point of view. The two hands are mirror images of one another, operating in an environment that repeats consistently from left to right. So either a chord or a visual pattern that is fingered 1-3-5 in the right hand, will be executed 5-3-1 in the left hand; a similar mirror transformation. This would not explain why the right hand/treble clef is faster in left-handed individuals, but provides one hypothesis for the difference between hands.

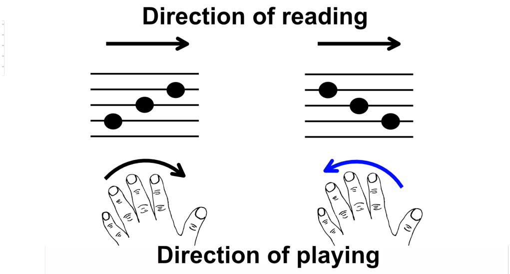

Figure 11. Playing/Reading incongruence in falling triads

In terms of decoding the instructions, several factors may be at work. We expect a two-fold effect of multivalent notation, both in obscuring the actual requirements at the point of execution, and in obstructing the long-term pattern learning that we know is a pre-requisite for good sight-reading. Disentangling these effects from one another, and from middle-term recency effects of one key signature on another is a complex but fascinating challenge. Pianists represent a population with thousands of hours of practice at a very complex and interlocking set of task-mappings, which is unmatched in normal experimental settings.

The effect of last-seen congruence reported in section III.D is an important finding. Apparently the clef of the last sighting of the a visual configuration (either in the same clef or the other clef) was more significant than whether that stimulus had been rehearsed in the previous block, where it would have been shown four times, twice in each clef. This suggests that beyond the small ‘task-switch’ effect, the switch might in itself have the capacity to disrupt recent learning.

In terms of longer term learning, there are two main lenses we might apply. The ’cellists (including one who was also left-handed), on being informed that they read more fluently in the right hand expressed little surprise, stating that they had learned the Treble clef first as children, and had always felt comfortable reading it. If early experience of a set of task rules forms the basis for music-reading, with other mappings either overlaid in conflict, or extended by some further transformation, we might also be able to explain the overwhelming advantage observed for the central key of C major/A minor, which most piano teachers introduce first.

On the other hand, accumulated exposure may play a substantial role in shaping long term learning. The treble clef effect may simply reflect the fact that the timbral qualities of the instrument mean there are usually more notes to play in the right hand than the left, those notes are somewhat less likely to form patterns that can be taken in at a glance, and consequently treble clef reading is more practiced. Whether the same can be said of the central key C major/A minor is more questionable, as any young learner who has spent many hours playing the Moonlight Sonata will attest.

There is some indication in the data that some keys may be read as ‘extensions’ of other related keys. Some 50% of all the single-note errors in the first and second experiments were semitone errors at the 4th or 7th degree of the scale, i.e. forgetting the last flat or sharp of the key signature. This seemed to be irrespective of the immediately preceding key signature. In addition, some of the diatonic graphs showed signs of skew towards a particular reading strategy. When chord IV appears to be the most fluently read in the key of 1#, for example, we may suppose that the key is essentially being read as “Cmajor with an extra sharp”. Additionally some participants’ individual graphs in 3b show an advantage for chords vi and iii (when 3 is usually the most disadvantaged chord after chord 7) which strongly suggests that 3b is being read preferentially as C minor, rather than as Eb major. These ‘extension’ strategies appear to be quite individual, however, and may reduce or disappear in experts; more data from balanced experiments is required to investigate this effect.

Key signatures may interfere with each other in a number of ways. One might imagine that A major (3#) followed by A flat major (4b) might present a particular set of difficulties, as every single note mapping will be altered. On the other hand, the diatonic chords remain in the same locations on the musical staff, which may be an advantage. In terms of pure spatial congruence, 3b is the mirror inverse of 3#, and playing these one after the other might also be expected to cause a particular pattern of disruption between the hands. A further possibility, that the major and minor key signatures of the same keynote either enable or disrupt one another could also be considered. None of these questions provide a direct hypothesis for why 3# should show longer average response times than 3b, but provide plausible avenues of enquiry which form the basis of the third experiment.

V. CONCLUSION

The multivalent nature of music notation, as discussed in the introduction, may provide a reasonable hypothesis for the particular difficulty of sight-reading in two clefs and in many keys. This experiment was set up to encourage 'chunking' of three-group notes into patterns that are amongst the most common in the literature, and to observe how better pattern recognition of these chunks may be acquired.

Further work aims to collect more data from expert participants to contrast with the moderate group of the second experiment, and to complete the key signature comparison design of the third experiment. Further study of the ways in which individuals select or extend core key signature mappings, and of the learning that takes place during the experiment is also planned.

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.

REFERENCES

Fourie, E. (2004). The processing of music notation: some implications for piano sight-reading. *Journal of the Musical Arts in Africa.*, *1*, 1–23.

Goolsby Thomas W. (1994). Eye Movement in Music Reading: Effects of Reading Ability, Notational Complexity, and Encounters. *Music Perception: An Interdisciplinary Journal*, *12*(1), 77–96.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*(3), 134–140. http://doi.org/10.1016/S1364-6613(03)00028-7

Sloboda, J. (1974). The Eye-Hand Span-An Approach to the Study of Sight Reading. *Psychology of Music*, *2*(2), 4–10. http://doi.org/10.1177/030573567422001

Vandierendonck, A., Liefooghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin*, *136*(4), 601–626. http://doi.org/10.1037/a0019791

Wood, M (2016). Visual Confusion in Piano Notation. In *Proceedings of the Second International Conference on Technologies for Music Notation and Representation* (In press)